University of Verona

Department of Diagnostics and Public Health

PhD school of Life and Health Sciences

PhD in Inflammation, Immunity and Cancer

XXIX CYCLE

The link between centrosome defects and cancer unveiled by *CROCC* deficiency in rhabdoid colorectal cancer

SSD MED/08

Coordinatore: Ch.ma Prof.ssa Gabriela Constantin

Tutor: Dott.ssa Erminia Manfrin

Dottorando: Dott Pietro Parcesepe

1 Abstract

The century-old relationship between centrosomes and cancer has remained a still unresolved question. Here we describe somatic mutations in *CROCC* gene, the major structural component of centrosome cohesion apparatus by sequencing the whole exome from two patients with a rare and lethal tumor referred to as rhabdoid colorectal cancer. We found that impaired *CROCC* activity at the "1p36.13" locus either caused by *de novo* somatic mutations or by copy number loss, results in aberrant centrosome phenotypes and unique catastrophic mitotic forms not recurrent in extrarenal rhabdoid or classical colorectal cancers. Our results demonstrate that reduced *CROCC* dosage, disrupts bipolar mitotic spindle architecture causing tetraploid DNA segregation errors and rhabdoid-like phenotype *in vitro*. In contrast, its restoration in an metastatic model harboring 1p36.13 deletion, arrests growth, corrects tetraploidy and centrosome segregation errors functioning as a biological barrier against replication stress and tetraploidy, two hallmarks of chromosomal instability. Our work reveals that impairing *CROCC* activity underlie rhabdoid tumors providing a link between centrosome genetic defects and human cancer, which may inform new therapeutic approaches to limit lethal phenotypes.

2. Introduction

2.1 Centrosome

2.1.1 Composition: centrioles and PMC

Centrioles constitute the core of the centrosome. Centrioles are cylindrical structures that are ~450 nm in height and ~250 nm in diameter, and characterized by a radial arrangement of nine peripheral microtubule triplets, as well as proximaldistal polarity along the cylinder. Centrioles are critical for recruiting the surrounding pericentriolar material (PCM). Although the PCM appears largely amorphous when observed by electron microscopy, super-resolution light microscopy reveals concentric layers of organization around the centriole. During interphase, an inner layer of PCM proteins is present next to centriolar microtubules, and notably contains the γ-tubulin ring complex, which is fundamental for microtubule nucleation. PCM architecture changes towards mitosis, with an expansion of the inner layer and the addition of further components, together resulting in a mature centrosome with maximal microtubule organizing centre (MTOC) activity. Whereas centrioles are crucial for PCM recruitment, the PCM reciprocally contributes to centriole biogenesis. [1]

2.1.2 Function

In most animal cells, centrosomes are a major source of spindle microtubules, and they are absolutely essential for cell division in several early embryonic systems.

It is clear, however, that centrosomes are not absolutely essential for division in many cell types. When centrosomes are absent, either naturally (as in higher plants or in the female germ cells of many animal species) or due to experimental manipulation, bipolar spindles can form in the vicinity of chromosomes through a centrosome-independent pathway that involves the small GTPase Ran and the action of microtubule motors and microtubule-bundling proteins. This pathway also presumably explains the surprising finding that Drosophila mutants lacking the centriole duplication protein DSas-4 appear to proceed normally through most of development, provided that a maternally supplied pool of DSas-4 is initially present to allow centrosome formation during the earliest stages of embryogenesis. In these mutants, centrioles and centrosomes are undetectable in adult cells, yet adults appear morphologically normal and eclose with nearnormal timing at near-

normal Mendelian ratios. This is in stark contrast to the rare larvae that develop to adulthood in many mutants that show dramatic defects in cell division. Thus, although spindle assembly is slowed in fly cells that lack centrosomes, flies appear to proceed through most of development relatively normally using only the centrosome-independent pathway of spindle assembly. Perhaps this reflects the fact that Drosophila cells have only four chromosomes to segregate, and organisms with larger numbers of chromosomes may depend on the greater efficiency of spindle assembly afforded by centrosomes—an intriguing notion that remains to be experimentally tested. Interestingly, Drosophila mutants that lack centrioles nevertheless die soon after they eclose. Rather than reflecting the absence of centrosomes, however, this death appears to result from the lack of cilia that are essential for the function of certain mechano- and chemosensory neurons. These observations support the view that centrioles may have originally acquired the ability to form centrosomes not to increase the efficiency of cell division but rather to ensure that the centrioles associate with the spindle poles and are thereby equally partitioned between the two daughter cells. Although many somatic cells can clearly divide without centrosomes, there is compelling evidence that centrosomes are required for the efficient division of cells that split asymmetrically to produce two daughter cells of different fates. Although centrosomes are not essential for cell division in all cell types, there is evidence that they contribute to efficient cell-cycle progression at both the G1/S and the G2/M transitions. In response to experimentally induced perturbation of centrosomes, some vertebrate cells undergo a G1 arrest, prompting speculation that a specific checkpoint might monitor the functional integrity of the centrosome.

During S phase, a subpopulation of cyclin E associates with centrosomes and possibly contributes to the regulation of S phase entry, and several cell-cycle regulatory proteins are concentrated at centrosomes and spindle poles during mitosis. These observations have led to the proposal that centrosomes might function as “scaffolds” to promote interactions between various regulatory components during the cell cycle. At the G2/M transition, the key mitotic kinases Cdk1/Cyclin B, Aurora-A, and Polo family members all accumulate at centrosomes, and the mitotic activation of Cdk1 is first detected at centrosomes. There are also several reports indicating that components of the DNA-damage checkpoint are concentrated at centrosomes. [2]

2.1.3 Duplication

Proteins that have critical roles in centrosome duplication cycle have been discovered more recently, setting the stage for a molecular interrogation of the relationship between centrosomes and cancer. Most proliferating cells are born with two centrioles that are loosely connected via a

proteinaceous linker. These two centrioles are usually close to one another in the G1 phase of the cell cycle, effectively constituting a single MTOC. Typically around the G1/S transition, one procentriole begins to assemble orthogonal to the proximal end of the mother centriole and the daughter centriole. The two procentrioles then elongate during the remainder of the cell cycle while remaining engaged with their neighbouring centriole. The spatial relationship of the two centrosomes changes at the G2/M transition. First, the proteinaceous linker connecting them is removed following activation of the serine/threonine kinase NEK2, in a step referred to as centrosome disjunction. Second, the two centrosomes separate along the nuclear envelope before nuclear envelope breakdown, in a step dubbed centrosome separation. This process is thought to be driven principally by kinesin 5, a tetrameric plus-end-directed motor that pushes apart overlapping microtubules located between the centrosomes. Thereafter, the two separated MTOCs direct the assembly of the bipolar spindle, which ensures faithful segregation of the genetic material to daughter cells. During mitosis, the centriole and procentriole disengage from one another within each centrosome, so that each daughter cell inherits two centriolar cylinders, thus completing the duplication cycle. [1]

2.1.4 Centrosome and cancer

Centrosome abnormalities, usually increased numbers, are common in human tumours and experimentally induced tetraploid cells from extra centrosomes, can be critical for aneuploidy and metastatic potential, however, the key factor underlying centrosome anomalies in tumorigenesis remains unclear. [3]

A signature feature of cancer is uncontrolled progression through the cell cycle. Therefore, understanding how the centrosome duplication cycle is coupled to the cell cycle is crucial for evaluating how centrosomes may participate to control proliferation. Core components of the cell cycle machinery regulate the centrosome duplication cycle, with two particularly important coupling points at the G1/S and G2/M transitions.

Furthermore it has been ever clear that cancer cells frequently exhibit aberrant mitotic figures, an observation that fuelled the postulate that extra centrosomes contribute to genome instability. Experiments with antibodies directed against centriolar components established the presence of centrosome aberrations, including excess centrioles, longer centrioles, excess PCM, as well as changes in centrosome shape and size, occurring both in cells from solid tumours and in haematological malignancies. Such aberrant figures can be present in pre-invasive lesions, and their

frequency often correlates with tumour progression raising the possibility that centrosome aberrations promote cancer progression.

The extent of extra centrosomes correlates with that of Chromosomal instability (CIN) in many human cancers. CIN that refers to the propensity of cell populations to change their chromosome complement over time, is a hallmark of cancer

Yet, a causal relationship between centrosome abnormalities and cancer has been difficult to establish. The deregulation of several oncogenes and tumor suppressor genes is well known to affect the number of centrosomes.

In this scenario, the rare and extremely aggressive entity referred to as rhabdoid colorectal cancer (RC), with a median survival of 6 months, remains uncharacterized. [1-2]

2.2 Rhabdoid colorectal cancer

2.2.1 Generality

Rhabdoid cancer is an extremely rare and highly aggressive tumor type with no proven effective treatment.[4]

The first reported case of rhabdoid tumor was a malignant pediatric renal neoplasia, a variant of Wilms tumor, with histologic features of rhabdomyoid-like appearance[5.] Malignant rhabdoid kidney tumors are considered a distinct clinicopathologic entity based on morphologic findings and a high metastatic potential. Tumors with similar features have been reported in several other organs such as esophagus, stomach, small intestine, heart, skin, and colon rectum. In this scenario, the rare and extremely aggressive entity referred to as rhabdoid colorectal cancer (RC), with a median survival of 6 months, remains uncharacterized. To date, only 23 cases of colorectal adenocarcinoma with a rhabdoid dedifferentiated phenotype have been reported, indicating that it is extremely rare[**4,6-23**] (TABLE 1). RCs occur in adults (range, 31-87 years) suggesting a tendency to develop in older patients and are morphologically indistinguishable from extrarenal rhabdoid tumors (ERTs) typically affecting infants and young adults. A common features of extrarenal rhabdoid tumors is the presence of metastasis at diagnosis and patient poor prognosis[6]. Clinically, the majority of colorectal rhabdoid cancers arise in cecum and transverse colon, and can display either a “pure” rhabdoid morphology when the rhabdoid features are the only identifiable phenotype or more often “composite” when the rhabdoid phenotype is mixed with another type of identifiable neoplasm. Histologically, the rhabdoid phenotype is characterized by the presence of pleomorphic cells with large, eccentric nuclei, prominent nucleoli, abundant and eosinophilic cytoplasm, paranuclear inclusions of intermediate filaments, and abundant mitotic figures. Cytokeratin and vimentin are

frequently found on immunochemistry[21]. It is equally distributed between both sexes with 13 males and 10 females. Almost all patients presented with abdominal symptoms including abdominal pain, abdominal mass, and gastrointestinal bleed. This is quite an unusual presentation of colon cancer that tends to have more occult presentation and could be a reflection of the aggressive nature of this type of tumor.

2.2.2 Molecular mechanisms

Previous efforts to gain insight into the molecular mechanisms underlying rhabdoid phenotype are limited to pediatric malignancies, which are characterized by low mutation load and a inactivating mutation and/or deletion of the chromatin remodelling gene *SMARCB1 (INI1)* residing at chromosome 22q11.2.

Infact Rhabdoid tumors are typically diagnosed in infants and children, but they can occur at any age. Rhabdoid tumors have been described in virtually every anatomic site, including the brain, soft tissue, lungs, ovaries, and liver, but they most frequently originate in the kidneys and brain. Rhabdoid tumors originating in any location receive the generic name of malignant rhabdoid tumors (MRT). However, when MRT arise in the kidney they are specifically called rhabdoid tumors of the kidney (RTK), and when they arise in the brain, they are called atypical teratoid/rhabdoid tumors (AT/RT). The name AT/RT was conceived by Rorke and colleagues when rhabdoid tumors of the central nervous system were defined as a new entity. It reflects the ‘‘unusual combination of mixed cellular elements similar to but not typical of teratomas’’ observed in that group of tumors. Currently, rhabdoid tumors originating at any anatomic location are recognized as the same tumor type, with similar morphology, biology, and clinical behavior. The cell origin of MRT remains unknown. In a report based on a microarray gene expression study, Sredeni and collegues suggested that rhabdoid tumors may arise through a developmental arrest of neural crest stem cells[24]

Treating AT/RT requires surgically resecting as much of the tumor as possible while preserving neurologic function, followed by an intensive multimodal regimen. Chemotherapy may include vincristine, cisplatin, cyclophosphamide, etoposide, actinomycin D, and temozolomide. The addition of intrathecal chemotherapy and radiation therapy depends on the child’s age and the extent of the disease . This regimen has significantly improved patients’ survival rates. Regrettably, despite the progress, the treatments’ toxic side effects are substantial, and most patients still rapidly succumb to their diseases. [25]

Regardless of their site of origin, the vast majority of MRT demonstrate abnormalities in chromosome 22. [26]

SMARCB1, stands for SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily b, member 1. The SWI/SNF chromatin-remodeling complex consists of 12–15 subunits and uses energy obtained from adenosine triphosphate (ATP) hydrolysis to remodel nucleosomes and modulate gene transcription. Mutations in genes encoding these subunits, structural abnormalities, or epigenetic modifications that lead to reduced or aberrant expression of members of the SWI/SNF complex have been reported in 20% of human cancers. The member of this complex encoded by the SMARCB1 gene is recruited to various chromatin regions, including gene promoters, that regulate cell cycle, growth, and differentiation The SMARCB1 abnormalities in rhabdoid tumors are characterized as somatically acquired biallelic inactivating truncating mutations within tumors with or without a predisposing germline mutation. This characteristic implicates SMARCB1 as a tumor suppressor gene, as defined by Knudson in the ‘‘two-hit model’’.[25]

Recently, a 2nd core element of the SWI/SNF chromatin-remodeling complex, the SMARCA4 gene, also known as BRG1 and located at 19p13.2, was found to be inactivated in rare cases of rhabdoid tumors that retained SMARCB1 expression, as detected by immunohistochemistry. [27]

Malignant rhabdoid tumors are polyphenotypic tumors that express markers of divergent differentiation. Rhabdoid tumors have a classic immune profile that shows diffuse expression of smooth muscle actin, epithelial membrane antigen, and vimentin. These markers are associated with variable expression of neuron-specific enolase, Leu7, and S100; absence of expression of muscle markers such as desmin and myogenin; and loss of expression of SMARCB1. Documenting loss of SMARCB1 protein expression is particularly useful to characterize the tumor and should be included in the final diagnosis. [25]

It has been estimated that up to one third of patients with rhabdoid tumors harbor SMARCB1 germline inactivating mutations. Although most of these mutations seem to occur de novo, familial cases have been reported in which an inherited constitutional SMARCB1 mutation of 1 allele predisposes a patient to developing a rhabdoid tumor. This condition is known as rhabdoid tumor predisposition syndrome (RTPS). [28]

By contrast, information on molecular anomalies of RCs is limited to rare cases previously described by (Pancione and Remo) wich showed microsatellite instability (MSI), *BRAF*^{V600E} mutation and CpG island methylator phenotype (CIMP). They analyzed promoter DNA methylation of subset of genes highly specific to characterize the CIMP status (*NEUROG1*, *IGF2*, *RUNX3*, *SOCS1*, including *MLH1*). From these analyses, they found *MLH1* promoter hypermethylation, which—associated with negative MLH1 immunostaining—confirms MSI-H.

In addition, aberrant DNA promoter methylation was identified in multiple loci: *NEUROG1*, *IGF2*, *RUNX3*, and *SOCS1*, suggesting the presence of a CIMP-positive tumor (data not shown). Altogether, these data implyed that rhabdoid features were closely correlated with presence of MSI, CIMP+ tumor, and BRAFV600E mutation. No driver mutations-associated phenotypes have been reported in any of these cases.

3. Material and methods

3.1 Research Subjects and Rhabdoid Tumors

Seven cases of primary rhabdoid tumour arising in colorectum (RC) and matched normal samples were recruited from different medical institutions: a) Konkuk University School of Medicine, Seoul University, Seoul, Korea; b) Hospital Santariskiu Clinics, National Affiliate of Vilnius University Hospital Santariskiu Clinics, Vilnius, Lithuania. c) Mater Salutis" Hospital, Legnago, Verona, Italy**.** The discovery samples included paraffin embedded specimens of rhabdoid tumour component and matched non-neoplastic mucosa from two patients with primary rhabdoid tumour arising in colorectum (RC), who were recruited at the "Mater Salutis" Hospital, Legnago, Verona and "G. Rummo" Hospital, and Benevento, Italy.[6,17,19, 20, 24] To be included in the study a centralized revision of all cases was performed, the hematoxylin-eosin–stained (H&E) glass slides were independently reviewed by pathologists to confirm the diagnosis and to look for the presence of strict rhabdoid morphology as previous described. Five out of the seven cases of the cohort have been reported previously. [6,17,19,20,29] Two cases have not been reported before (VI, VII). Clinicopathologic features of RCs are summarized in (**Table 1)**. Additional seven cases of "extrarenal rhabdoid tumors" affecting central nervous system of patients between 2 months and 19 years of age, diagnosed between 2007-2015, were recruited from the "Azienda Ospedaliera Universitaria Integrata", Verona, Italy. The pediatric rhabdoid tumors were indicated as Rhabdoid of infant (RI) through the study. The main clinicopathologic features of the 7 pediatric rhabdoid tumors are summarized in (**Table 2)**. For all cases of rhabdoid tumors, matched normal-tumor pairs consisting of formalin fixed paraffin-embedded (FFPE) tissue blocks, at the time of the initial surgery, were collected.

Table 1. Rhabdoid colorectal cancers: clinicopathological, immunohistochemical and molecular characteristics of the 7 cases

Clinicopathological features

Case	Age (years)	Sex	Location	Size (cm)	pT	Metastasis	Follow up	Status	Reference
RC 1*	73	Female	Cecum	10x8	4	Lymph nodes	6 months	Dead	#
RC 2*	71	Male	Cecum	10x10	4	Lymph nodes	8 months	Dead	#
RC 3	62	Male	Sigma	4.5x4	3	Lymph nodes	36 months	Alive	#
RC 4	83	Female	Rectum	6.5x4.3	4	Liver, Lung	1 month	Dead	#
RC 5	49	Male	Sigma	7x7	4	Lymph nodes	7 months	Dead	#
RC 6	63	Male	Left colon	6 x6	3	Lymph nodes	1 month	Dead	none
RC 7	71	Female	Cecum	8x11	4	Lymph nodes	8 months	Dead	none

Immunohistochemical analysis

Case	CK20	CDX2	Vimentin	Mhl1	Pms2	Msh2	Mlh6	p53	SmarcB1/ Ini1
RC 1*	-	-	+	-	-	+	+	++	neg.
RC 2*	-	-	+	-	-	+	+	++	++
RC 3	-	+	+	+	+	+	+	-	++
RC 4	-	-	+	+	+	+	+	-	++
RC 5	-	-	+	+	+	+	+	None	++
RC 6	-	-	+	+	+	+	+	None	neg.
RC 7	-	-	+	-	-	+	+	++	++

Molecular analysis

Cases	MSI	CIMP	BRAF	*KRAS*	APC	TP53
RC 1*	MSI-H	pos (6/6)	MUT	WT	WT	WT
RC 2*	MSI-H	pos (5/6)	MUT	WT	WT	WT
RC 3	MSS	pos (3/6)	WT	WT	MUT	WT
RC 4	MSS	Neg(1/6)	MUT	WT	WT	MUT
RC 5	MSS	pos (5/6)	MUT	WT	WT	WT
RC 6	MSS	pos (4/6)	MUT	MUT	MUT	MUT
RC 7	MSI-H	Pos (5/6)	MUT	WT	WT	MUT

Notes and abbreviations: *, cases subjected to exome sequencing; pT, pathological stage, where T3 is infiltration of the intestinal wall, and T4 is extension to perivisceral fat; +, positive immunostaining; -, negative immunostaining.

Abbreviation: Tumors with at least three methylated loci of the following CIMP loci (*RUNX3, IGF2, SOCS1, NEUROG1, CDKN2A* and h*MLH1*) were classified as CpG island methylator phenotype (CIMP)-positive and the remaining cases as CIMP-negative.

Table 2. Infants/young adults rhabdoid cancers of the central nervous system: clinicopathological and immunohistochemical features of the 7 cases.

Case	Age (years)	Sex	Location	INI1 (IHC)	CROCC (IHC)	Follow up (months)	Status
RI 1	2	F	CNS	-	-	n.a.	n.a.
RI 2	2 months	M	CNS	-	Single, large centriol	3	Dead
RI 3	4	M	CNS	+	Single, large centriol	33	Alive
RI 4	5	M	CNS	-	Single, large centriol	21	Dead
RI 5	10	F	CNS	+	Single, large centriol	92	Alive
RI 6	6	M	CNS	-	Single, large centriol	n.a.	n.a.
RI 7	19	F	CNS	-	Single, large centriol	46	Alive

CNS, central nervous system; AT/RT, Atypical teratoid rhabdoid tumor
n.a., Not Available

3.2 FFPE DNA extraction.

The tumoral area with rhabdoid morphology was identified by H&E staining, whereas, manual macrodissection was performed on the FFPE blocks using a scalpel. Briefly, samples were then incubated at 90 °C to remove DNA crosslinks, and extraction was performed using Qiagen's QIAamp DNA FFPE Tissue Kit according manufacturer's instructions. DNA concentration was determined with both spectrophotometric Nanodrop 1000 (Thermo fisher, Wilmington, DE, USA) and fluorometric Picogreen dsDNA kit (Invitrogen, Inchinnan, UK) methods. On average, the concentration measured with picogreen was half the concentration estimated with Nanodrop.

3.3 Whole-Exome Sequencing

Isolated genomic DNA "1.3 µg" was sheared to 100-450 bp on a Covaris S220 instrument (Covaris, Woburn, Ma, USA). Distribution of fragments size was verified with an Agilent 2100 Bioanalyzer using a DNA 1000 assay (Agilent, Santa Clara, CA, USA). Sheared DNA was subjected to Illumina paired-end DNA library preparation using the TruSeq DNA Sample preparation kit (Illumina, San Diego, CA, USA) and enrichment for whole exome was carried out according to the TruSeq Exome Enrichment Guide (Illumina, San Diego, CA, USA). Briefly, fragmented DNA was end-repaired and adenylated before the ligation of an indexing adapter. After the subsequent PCR amplification, the quality of the library was evaluated with an Agilent High Sensitivity DNA assay (Agilent, Santa Clara, CA, USA) and then quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo fisher, Wilmington, DE, USA) on a NanoDrop 3300 fluorospectrometer (Thermo Scientific, Wilmington, DE, USA). Exome capture was performed using 500 ng of library as input material. Next, half biotinylated bait-based hybridizations were performed, each followed by Streptavidin Magnetic Beads binding, a washing step and an elution step. Libraries were amplified with 10 PCR amplification cycles. The quality of the whole exome library was checked with an Agilent High Sensitivity DNA assay and quantified by qPCR on a Stratagene MX3000P (Agilent, Santa Clara, CA, USA) using Kapa Library Quant kit (Kapa Biosystems, Boston, MA, USA). Whole exome library was sequenced with an Illumina HiSeq 1000 sequencer (Illumina, San Diego, CA, USA) and 100-bp x 2 paired-end sequences were generated. The sequences were pre-processed by removing reads with a percentage of undetermined bases (N) higher than 10% of the length of the read and more than 50 bases with a quality lower or equal to 7. Adapters were clipped using Scythe v0.980 (https://github.com/vsbuffalo/scythe) and 3′ ends with a quality score lower than 20 over a window of 10 bases were trimmed using Sickle v0.940 (https://github.com/vsbuffalo/sickle), eventually entirely removing the fragment if the final length of one of the reads was lower than 50 bp. Filtered reads were aligned against the reference hg19 genome using BWA 0.6.2 using default parameters.[30,31] Duplicates were marked with Picard tools and reads were realigned in correspondence of indels and recalibrated using GATK 2.6-5.[32,33] Variants were called with Mutect v1.1.4 using default parameters and with GATK 2.6-5 UnifiedGenotyper module with "-glm BOTH" parameter.[34] Functional annotation of mutations was performed with Annovar and cancer driver analysis with CRAVAT.[31]

3.4 Tumour purity, copy number and structural variations analysis

The purity of the tumor samples was estimated using SNP array data with the Allele-specific copy number analysis of tumors (ASCAT) algorithm as reported.[35] SNP arrays were scanned and data

was processed using the Genotyping module (v1.8.4) in Genome studio v2010.3 (Illumina, San Diego, CA, USA) to calculate B-allele frequencies (BAF) and logR values. GenoCN43 and GAP17 were used to call somatic regions of copy number change – gain, loss or copy neutral LOH. Recurrent regions of copy number change were determined and genes within these regions were extracted using ENSEMBL v70 annotations. Depending on the read pair types supporting an aberration or the associated of copy number events each structural variant was classified as: deletion, duplication, tandem duplication, foldback inversion, amplified inversion, inversion, intrachromosomal or translocation. Essentially, the type of rearrangement is initially inferred from the orientation information of discordant read pairs, soft clipping clusters and assembled contigs which span the breakpoints. Structural variants with breakpoints that flanked a copy number segment of loss were annotated as deletions. Duplications and inversions associated with increases in copy number enabled the characterization of tandem duplications and amplified or foldback inversions. Events within the same chromosome which linked the ends of copy number segments of similar copy number levels were identified. Events were then annotated if they were within 100 kb of a centromere or telomere and genes which were affected by breakpoints were annotated using ENSEMBL v70. Somatic copy number variation, structural variants and copy number data were visualized using circus.[36]

3.5 Splice-site prediction tool and driver genes analysis

To predict the splicing mutations that affected donor and acceptor splice sites and to evaluate the efficiencies of physiological splicing sites in mutant genomic sequence Mutation Taster tool was used. To further evaluate the disease-causing potential on both exonic and nonexonic variants, we also used SpliceFinder which is a method for rapid functional prediction of splicing variants starting from a large set of somatic mutations obtained by whole-exome sequencing analysis.[13] The SpliceFinder methodology is a bioinformatics integrated procedure based on two public functional annotation tools for HTS analysis, ANNOVAR and MutationTaster and two canonical splice-site prediction software programs for single splicing analysis, SSPNN (http://www.fruitfly.org/seq_tools/splice.html and NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/.[37-39] The candidate variant was accepted as putative splicing mutation and sequenced by Sanger method whereby the prediction was confirmed using both programs. To

identify putative driver mutations that are required for the cancer phenotype, the computational tool, DrGaP (driver genes and pathways) was applied to the common mutated genes from cancer genome-sequencing. This tool incorporates statistical approaches and several auxiliary bioinformatics tools for better driver gene identification.[40]

3.6 Sanger sequencing validation

We used polymerase chain reaction (PCR) amplification and direct DNA sequencing to validate candidate variants in *CROCC* gene identified by WES (**Table 3)**. Because of the sensitivity of Sanger, sequence variants that were reported in less than 20% of the reads could not be included in this validation phase. The purified products were subsequently sequenced with the use of the BigDye Terminator, version 3.1, Cycle Sequencing Kit (Life technology, Carlsbad, CA, USA). Data were analyzed with BioEdit Sequence Alignment Editor 7.0.5.3.

Table 3. reverse transcription and sequencing PCR primers

CROCC primers		
Gene	Sense	Antisense
RTPCR-CROCC-II	5’ACATGACCGCCAAGTACTCC3’	5’CCTCCCGTAGCTGTTGTAGC3’
RTPCR-CROCC-I	5’CTGAGGGACACAGAGCACAG3’	5’TCCTCACTCAGAGCCTGGTT3’
Seq-CROCC	5’AAATTGGAGGAGACGGCTTT3’	5’CAGCTTCTGCTCCTTGTCCT3’
RAC1	5′-GCCAATGTTATGGTAGAT-3′	5′-GACTCACAAGGGAAAAGC-3′
CDX2	5’ AAAGTGAGCTGGCTGCCACACTTG 3’	5’ TCCATCAGTAGATGCTGTTCGTGG 3’
KRT20	5’CTGAATAAGGTCTTTGATGACC 3’	5’ATGCTTGTGTAGGCCATCGA 3’
CDH1	5’GAACGCATTGCCACATACAC 3’	5’ATTCGGGCTTGTTGTCATTC 3’
LAMA5	5’AAGATGGCGAAGCGGCTCT 3’	5’ TTGAAGTAGGGCGGGTGCAG 3’
KRT18	5’ TTTCCCAGATCATGGAGGAG 3’	5’AGCCCATGAGGTTTTTCTGA3’
VIM	5’TGACAATGCGTCTCTGGCAC 3’	5’CCTGGATTTCCTCTTCGTGG 3’
18S	5’GGGAGCCTGAGAAACGGC 3’	5’GGGTCGGGAGTGGGTAATTT 3’
ACTB	GCTCACCATGGATGATGATATCGC	ATAGGAATCCTTCTGACCCATGCC
ACTG2	5’-CCGCCCTAGACATCAGGGT-3’	5’-TCTTCTGGTGCTACTCGAAGC-3’
SMARCB1	5’TCCGTATGTTCCGAGGTTC3’	5’CTTCCACTTCCGAGGCTTT3’
CEP152	5'-CCATGTCAGCTTGTTCTTCC-3'	5'-CACCTTTCTCTTCTCCTGCT-3'
CEP192	5’CCCAACGACCTAATGATGTTCA3’	5’GCTCCCAAGTCGCTTGTAGATT3’
CDKN2B	5’TACAGGAGTCTCCGTTGGC3’	5’GTGAGAGTGGCAGGGTCTG3’
CDKN2A	5’CCACCCTGGCTCTGACCAT3’	5’GCCACTCGGGCGCTG3’
CDH2	5'-GGCATAGTCTATGGAGAAGT-3'	5'-GCTGTTGTCAGAAGTCTCTC-3'

3.7 Public Database as a Reference and Sequencing data

We used gene expression data estimated from whole exome sequencing (WES), whole transcriptome sequencing (WTS), to select possible functional genetic changes in our study. Because RC is a rare disease and obtaining control samples is not easy, we used a public database as a reference. Preprocessed level 2 somatic mutation data, Affymetrix SNP 6.0, were downloaded (n=631 CRC samples) by using the most recent update of the raw sequencing TCGA data set.[41] The logR ratios and allelic differences and copy number were estimated by the Affymetrix Genotyping Console.[41,42] Ploidy was estimated by calculating the weighted median copy number across all copy number segments, with weights equal to the segment length. Copy number segments of loss and gain were defined relative to the ploidy status of each sample by subtracting the ploidy estimate from the estimated copy number of the segment for which SNP 6.0 copy number data was also available. For RNA-seq data, enrichment analysis was performed using the R package GOseq to correct for bias due to transcript length, considering only categories with at least ten annotated genes. We annotated all samples contained within it either as colorectal cancer (n = 382) or normal colon mucosa (n = 21) from TCGA data set, and we compared the RPKM value of specific genetic changes found in our analysis with normal colon expression values.[41] For copy number alteration (CNA) analysis of WES data, we used the Copy Number Analysis for Targeted Resequencing (CONTRA) tool and summarized the exon-level log2 fold changes of read depth between the normal and tumor samples into gene-level log2 fold changes.[43] The bioinformatics search of *CROCC* gene mutation across cancer types started from a large collection of publicly available human gene-expression array experiments downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo) and TCGA . For recurrence of *CROCC* somatic mutations, "mutation data were filtered for exons and at splice sites (±3 bp)", copy number changes, mRNA transcript levels across different tumour types and cancer cell lines we used datasets from TCGA (https://tcga-data.nci.nih.gov) and cBioPortal for Cancer Genomics (http://www.cbioportal.org). The colorectal, rhabdoid or neuroblastoma cancer cell lines studied were obtained from cBioPortal for Cancer Genomics. Colorectal cancer cells were classified as CIN+ or (aneuploid) or CIN- (non-aneuploid) using weighted GII .0.2. For the 60 colorectal cancer cell lines, the total number of putative somatic mutations ranged from 0 to 1308. Cell lines with a mutation prevalence of >25 per 10^6 bases were designated as hypermutated, as reported.[44,45] These latter, showed a modal chromosome copy

number of 2n, by contrast, nonhypermutated cells tended to exhibit unstable profiles with modal chromosome copy numbers ranging from 2n to 4n in cell lines. [44,45] To define the gene-array expression thresholds used to separate cancer cell lines harboring 1p36.13 deletion, from those retaining 1p36.13 locus supervised clustering analysis with FDR<0.01 based on spearman and ward linkage was used. The *P values* were adjusted for multiple testing with the Benjamini and Hochberg method.[46]

3.8 Gene ontology enrichment analysis

The enrichment analysis was performed with the TopGO package using the classic algorithm and Fisher's test with the same cutoffs described above.[44-46] Genes tested for differential expression were used as the background. To create enrichment treemaps, parent categories that had enriched children were first removed, and maps were then created with the Treemap package, color coding categories according to the combination of non-overlapping parent categories accounting for the largest proportion of plotted categories. All reported *P values* were calculated using the Benjamini-Hochberg method. To understand causal connections between diseases, genes and networks of upstream or downstream regulators the samples were subjected to Ingenuity Pathways Analysis (IPA) (Ingenuity Systems) which was used as a starting point for building biological networks.

3.9 *CROCC* target sequencing

A multigene panel was developed using AmpliSeq designer software v2.1 to investigate the coding sequences of *CROCC* (NM_014675.3) genes. Details on target regions of the panel are reported in (**Table 4)**. Suitability of extracted DNA from FFPE sections was evaluated and quantified as described above. To further verify DNA integrity a multiplex a multiplex PCR was performed according to previous studies. [47-49] Briefly, twenty nanograms of DNA were used for each multiplex PCR amplification. Then, PCR amplified fragments to build an adequate library for deep sequencing were successfully obtained. The mean read length was 112 base pairs and a mean coverage of 6290x was achieved, with 94.1% target bases covered more than 100x. A minimum coverage of 20x was obtained in all cases. The quality of the obtained libraries was evaluated by the Agilent 2100 Bioanalyzer on-chip electrophoresis (Agilent, Santa Clara CA, USA). Emulsion PCR to construct the libraries of clonal sequences was performed with the Ion OneTouch™ OT2 System (Life technology, Carlsbad, CA, USA). Sequencing was run on the Ion Personal Genome Machine (Life technology, Carlsbad, CA, USA) loaded with Ion 318 Chip v2 (. Data analysis, including

alignment to the hg19 human reference genome and variant calling, was done using the Torrent Suite Software v.5.0 (Thermo fisher scientific, Whaltam, MA, USA). Filtered variants were annotated using a custom pipeline based on vcflib (https://github.com/ekg/vcflib), SnpSift, the Variant Effect Predictor (VEP), software and NCBI RefSeq database. Additionally, alignments were visually verified with the Integrative Genomics Viewer (IGV) v2.3 to further confirm the presence of mutations identified by targeted sequencing.[49-51]

Table 4. *CROCC* custom panel used for the next-generation target sequencing.

Gene	Chr	Start	End	Gene	Chr	Start	End	Gene	Chr	Start	End
CROCC	chr1	17248461	17248565	CROCC	chr1	17271988	17272118	CROCC	chr1	17287230	17287347
CROCC	chr1	17248554	17248670	CROCC	chr1	17272134	17272224	CROCC	chr1	17287637	17287723
CROCC	chr1	17249085	17249217	CROCC	chr1	17272675	17272775	CROCC	chr1	17292117	17292245
CROCC	chr1	17249160	17249244	CROCC	chr1	17272738	17272848	CROCC	chr1	17292222	17292343
CROCC	chr1	17250720	17250829	CROCC	chr1	17273186	17273323	CROCC	chr1	17292322	17292437
CROCC	chr1	17250799	17250931	CROCC	chr1	17274660	17274791	CROCC	chr1	17292383	17292508
CROCC	chr1	17250873	17250992	CROCC	chr1	17274780	17274917	CROCC	chr1	17292504	17292599
CROCC	chr1	17250976	17251089	CROCC	chr1	17274933	17275069	CROCC	chr1	17292628	17292729
CROCC	chr1	17256284	17256403	CROCC	chr1	17275203	17275319	CROCC	chr1	17292742	17292867
CROCC	chr1	17256414	17256546	CROCC	chr1	17275297	17275411	CROCC	chr1	17292883	17293016
CROCC	chr1	17256533	17256633	CROCC	chr1	17275424	17275510	CROCC	chr1	17293026	17293116
CROCC	chr1	17256637	17256752	CROCC	chr1	17277375	17277491	CROCC	chr1	17294680	17294776
CROCC	chr1	17256905	17257024	CROCC	chr1	17277480	17277608	CROCC	chr1	17294770	17294880
CROCC	chr1	17257013	17257116	CROCC	chr1	17277597	17277687	CROCC	chr1	17294899	17294999
CROCC	chr1	17257674	17257770	CROCC	chr1	17279663	17279769	CROCC	chr1	17295543	17295679
CROCC	chr1	17257742	17257870	CROCC	chr1	17279774	17279898	CROCC	chr1	17295668	17295803
CROCC	chr1	17257818	17257943	CROCC	chr1	17279902	17280005	CROCC	chr1	17295792	17295910
CROCC	chr1	17263077	17263164	CROCC	chr1	17279976	17280095	CROCC	chr1	17296213	17296351
CROCC	chr1	17263141	17263233	CROCC	chr1	17280645	17280757	CROCC	chr1	17296329	17296458
CROCC	chr1	17263263	17263365	CROCC	chr1	17280746	17280866	CROCC	chr1	17296433	17296527

CROCC	chr1	17263352	17263433	CROCC	chr1	17280826	17280935	CROCC	chr1	17296665	17296798
CROCC	chr1	17264084	17264185	CROCC	chr1	17281146	17281235	CROCC	chr1	17296794	17296903
CROCC	chr1	17264174	17264285	CROCC	chr1	17281217	17281347	CROCC	chr1	17296879	17296990
CROCC	chr1	17264739	17264868	CROCC	chr1	17281374	17281456	CROCC	chr1	17296947	17297081
CROCC	chr1	17264859	17264949	CROCC	chr1	17281702	17281837	CROCC	chr1	17297117	17297245
CROCC	chr1	17264963	17265050	CROCC	chr1	17281854	17281992	CROCC	chr1	17297164	17297289
CROCC	chr1	17265348	17265463	CROCC	chr1	17282004	17282124	CROCC	chr1	17297842	17297965
CROCC	chr1	17265665	17265771	CROCC	chr1	17282412	17282530	CROCC	chr1	17297937	17298073
CROCC	chr1	17266338	17266447	CROCC	chr1	17282518	17282647	CROCC	chr1	17298176	17298280
CROCC	chr1	17266441	17266572	CROCC	chr1	17282591	17282712	CROCC	chr1	17298746	17298876
CROCC	chr1	17266584	17266695	CROCC	chr1	17284990	17285092	CROCC	chr1	17298849	17298930
CROCC	chr1	17270505	17270588	CROCC	chr1	17285081	17285215	CROCC	chr1	17298955	17299075
CROCC	chr1	17270780	17270893	CROCC	chr1	17285248	17285359				
CROCC	chr1	17271855	17271963	CROCC	chr1	17287132	17287241				

3.10 *RAS* and *BRAF^{V600E}* Mutation Analyses

RAS and BRAF^{V600E} mutations were investigated on DNA extracted from FFPE tumor tissues using real-time polymerase chain reaction (PCR) on RotorGene Q MDx instruments (Qiagen, Hilden, DE). Scorpions® and ARMS® technologies were employed to detect KRAS mutations at the codons 12 and 13 (exon2) and *BRAF^{V600E}* mutation at the codon 600 (V600E/Ec,D,K,R) against the background of wild type genomic DNA (Therascreen® RGQ PCR Kit) (Qiagen, Manchester UK).

The therascreen® NRAS Pyro® (Qiagen, Manchester UK) was used to determine additional mutations at the codons 61 (exon3) of the *NRAS* gene. While the RAS Extension Pyro® kit was used to detect mutations at the codons 59,61 (*only KRAS*) (exon3) and 117, 146 (Exon4) of the *KRA*S and *NRAS* genes, respectively according manifactures' istructions (Qiagen Manchester UK). The DNA segment of interest were amplified by PCR and sequenced separately. The distinct patterns in the Pyrogram® trace were distinguishable from the background obtained from wild-type samples.

3.11 Targeting sequencing Hotspot Panel.

DNA was extracted from FFPE tumor samples using the GeneRead DNA FFPE Kit (Qiagen, Hilden, DE). DNA quality and quantity were assessed using the Qubit Fluorometer (Invitrogen, Inchinnan, UK) and the 2200 Tape Station instrument (Agilent, Santa Clara CA, USA). Good quality genomic DNA was subjected to library preparation prior to sequencing. NGS analyses were performed on the Ion Torrent Personal Genome Machine (PGM, Life Technologies, Grand Island, USA). Tumor samples were tested using a commercially available library kit, the Ion AmpliSeq Cancer Hotspot Panel v.2 (CHP2) which allows for simultaneous amplification of 207 amplicons in hotspot areas of 50 oncogenes and tumor suppressor genes. Each amplicon library was generated starting from 10 ng of gDNA, as indicated by the manufacturer and the multiplexed PCR to amplify targets was performed in 20 cycles. Samples were barcoded with Ion Xpress Barcode Adaptors Kit (Life technologies, Carlsbad, CA, USA to allow for discrimination between samples within a NGS run. DNA Library quantification was performed using the PCR quantification kit and the Quantstudio 12k flex real time PCR system (Life technologies, Carlsbad, CA, USA, diluted in nuclease-free water to obtain a final concentration of 100 pM. The library was mixed with Ion Sphere Particles (ISPs) and the subsequent emulsion PCR and enrichment were performed using the Ion PGM™ Template OT2 200 Template Kit (Life technologies, Carlsbad, CA, USA. The quality of the emulsion PCRs was measured using the Qubit IonSphere Quality control kit (Life technologies, Carlsbad, CA, USA (Life Technologies). IonSphere were sequenced on Ion 316 Chip using the Ion PGM™ Sequencing 200 Kit (Life Technologies). Only sample sequences with at least a quality score of AQ20 (1 misaligned base per 100 bases) were considered for further analyses. Data obtained was processed initially using the Ion Torrent platform-specific pipeline software Torrent Suite to generate sequence reads, trim adapter sequences, filter, and remove poor signal and low quality reads. Mutations were called by the Torrent Variant Caller plug-in v 4.4.3.3 (Life technologies, Carlsbad, CA, USA. Subsequently, data were uploaded from Torrent Suite™

Software to Ion Reporter Software v 4.6 and analyzed with "*AmpliSeq CHPv2 single sample"* workflow.

3.12 Immunohistochemical staining and analysis.

Immunohistochemical analysis for CROCC was performed on tumor and matched normal tissue sections. Four µm-thick sections were obtained from FFPE blocks. They were deparaffinised with BOND DEWAX Solution (Leica biosystems, Newcastle, UK) placed in graded alcohol solutions, washed, and pretreated with the Epitope Retrieval Solution 2 (EDTA buffer pH8.8) at 98 °C for 20 minutes. After the washing steps, peroxidise blocking was carried out for 5 minutes using the Bond Polymer Refine Detection Kit DS9800 (Leica biosystems, Newcastle, UK). Then, the sections were incubated with with a rabbit primary anti-human CROCC polyclonal antibody, clone NBP1-80820, rabbit IgG, dilution 1:200, (ABCAM, Cambridge, UK) for 15 and secondary antibody for 8 minutes, respectively. Subsequently, they were incubated with polymer (8 min), revealed with DAB-chromogen (10 min), and stained with hematoxylin (10 min). Fallopian tube (FT) tissues were used to develop CROCC immunohistochemistry protocol as recommended by supplier (positive control). In normal cells CROCC staining consisted of one or up to two dot-like signals. Consistently, the expression pattern was stratified into three categories: a) Loss, less than 1 per cell b) normal, 1-2 signals per cell. c) Amplified; more than 2 signals per cell. Two hundred cells in triplicate sections per sample were blindly scored by three authors (A.R, E.M, M.P). CROCC staining at the centrosomes was also considered abnormal if showed: 1) a diameter greater than twice the diameter of centrosomes present in normal epithelium within the same section; 2) excesses of centrioles or longer centrioles; 3) changes in shape, size or perinuclear position. An antibody to the centrosome protein gamma-tubulin (clone TU-30, dilution 1:1000) was used to counterstain the centrosome (**Table 5)**. The minimal criterion for the identification of a gamma-tubulin-positive cell in the context of an abnormal staining pattern associated with putative centrosome dysfunction, was the detection of two or more punctate, dot-like immunoreactive signals, and/or robust diffuse staining, in the cytoplasm of individual tumor cells as previously described. [52,53] Thus, signal score was analyzed taking into account the cytosolic "non-centrosomal" and centrosomal localization of gamma-tubulin in each tissue section. Samples were also immunostained with the following markers: CDX2, CK20, TP53 and SMARCB1/INI1 and MMR as previously reported.[54]

Antibody	Company	Clone	Source	dilution
α tubulin	ABCAM	18251	rabbit polyclonal	1:100
β tubulin	COVANCE	TUJ1	mouse monoclonal	1:500
γ tubulin	ABCAM	TU-30	mouse monoclonal	1:100
CROCC	ABCAM	NBP1-80820	rabbit polyclonal	1:200
INI1	BD	BAF47	mouse monoclonal	1:50
γH2AX	ABCAM	p S139	rabbit polyclonal	1:500

Table 5. Antibodies employed and working conditions for immunohistochemistry and immunofluorescence analysis

3.13 Double-labeling immunofluorescence analysis.

FFPE tissue sections derived from RC were immunostained for CROCC and γ-tubulin by immunofluorescence as described previously.[55] Deparaffinized sections were incubated for 50 min in a freshly made PBS, serum, solution containing 1 mg/ml of sodium borohydride (Sigma-Aldrich, Saint-Louis, MO, USA) to decrease the autofluorescence, and subsequently they were rinsed extensively with PBS. For double-labeling immunofluorescence studies on histological sections, anti-human CROCC polyclonal antibody and a mouse anti-human γ-tubulin monoclonal antibody above described were employed. Secondary fluorochrome conjugated antibodies included anti-rabbit AlexaFluor 594 and anti-mouse AlexaFluor 488, respectively, and were diluted 1:1000 with PBS. Double staining was performed in two sequential sessions of immunofluorescence staining. 4',6-diamidino-2-phenylindole (DAPI) was used to label cell nuclei. Slides were cover-slipped using an aqueous-based mounting medium (Vector Laboratories, Cambridgeshire, UK).

3.14 High resolution melting (HRM) analysis

Loss of heterozygosity (LOH) analysis was done using fluorescent labeled forward primers, followed by fragmental analysis detection on Rotor-gene Q 5plex HRM (Qiagen, Hilden, DE). DNA from rhabdoid tumor and matched normal samples were analyzed for LOH at the *CROCC* locus (1p36.13) by using three markers D1S3391, D1S1443, D1S3669. Additional markers at the SMARCB1 locus (22q11.2) D22S301 and D22S345 were investigated.[56] PCR conditions for the multiplex PCR consisted: initial denaturation at 95°C for 15 min, followed by 19 cycles of 95°C 30 s, 55°C 90s, 72°C 1min, and final extension at 72°C 30 min. PCR reactions were performed in a 10 µl reaction mixture by using 50 ng template DNA, 0.2 mM dNTPs (Roche, Penzberg, DE), 0.4 U FastStart Taq DNA Polymerase (Roche, Penzberg, DE), 1x fluorescent dye LCGreen Plus (Idaho Technology, Salt Lake City, UT, USA), 2 mM MgCl2 and forward and reverse primers (0.5 mM each) for each gene segment. PCR conditions were optimized to temperatures between 52°C and 64°C for each segment. After 30 cycles of amplification, PCR products underwent an additional 1 min at 98°C and then 5 min at 40°C to promote heteroduplex formation. Each capillary was then transferred to the High Resolution Melter instrument (HR-1) (Idaho Technology, Salt Lake City, UT, USA) for high resolution melting and curve analysis. Samples were melted at 0.2°C/s ramp rate.[33] Melting profiles were analyzed with HR-1 software using fluorescence normalization, temperature shift and conversion to difference and derivative plots. To confirm the reliability of the HRM assay, selected paired non-neoplastic tissue samples showing or not a difference in melting profiles were purified and then sequenced in both directions using the Big Dye Terminator 1.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA). Sequencing reaction was performed on an automatic sequencer ABI PRISM 310 Genetic Analyzer and sequences were analyzed using BioEdit program.[57]

3.15 cDNA preparation and quantitative PCR

Total RNA was isolated from three 10-µm FFPE paired tumour and normal tissue sections per patient using RNeasy FFPE Kit (Qiagen, Hilden, DE) according manufactures' instructions. Paraffin was first removed by xylene extraction followed by ethanol wash. A DNase I treatment step was included to remove DNA from total nucleic acids. Reverse transcription was performed using the first-strand cDNA synthesis kit (General Electric, Buckinghamshire, UK). qPCR was performed in 96-well plates using pre-designed TaqMan probe/primers on a ABI 7900HT system (Life technology, Carlsbad, CA, USA). All reactions were performed in duplicate. The relative amount of mRNA was calculated using the comparative Ct method after normalization to GAPDH or rRNA 18S expression. For RT-qPCR analysis, a total cDNA amount corresponding to 10–50 ng

of starting RNA was used for each reaction. Fast SYBR Green Master Mix from Life Technologies and 10 μM for each primer pair were used. qPCR reactions were performed on an Applied Biosystems 7500 Real-Time PCR machine using the standard amplification protocol. The sequences of the specific primers for RT-PCR are listed in (**Table 3)**. RNA purity and quantity were measured on a Jenway Genova Plus spectrophotometer (Bibby Scientific, Staffordshire,UK).

3.16 Protein extraction and immunoblotting

The Qproteome FFPE Tissue Kit, (Qiagen, Hilden, DE) was used to isolate full-length proteins from FFPE tissues according to the manufactures' instructions. To lyse cultured cells and frozen tissues, RIPA buffer (150 mM NaCl, 1.0% Nonidet P-40 or Triton X-100, 0.5% sodium deoxy-cholate, 0.1% SDS and 50 mM Tris, pH 8.0) supplemented with protease inhibitor cocktail (Sigma-Aldrich, Saint-Louis, MO,USA) were used. Next lysates were sonicated using the Bioruptor Sonication System (Diagenode, Liège, BE) and then centrifuged at 13,000*g* for 15 min. Protein quantification was performed using the Bradford protein assay (Bio-Rad, Hercules, CA, USA) following the manufacturer's instructions. Protein extracts (50-80 μg per sample) were run on a precast NuPAGE 4–12% Bis-Tris Gel (Life technology, Carlsbad, CA, USA) transferred to a nitrocellulose membrane and blocked in TBST (50 mM Tris, pH 7.5, 150 mM NaCl and 0.1% Tween-20) and 5% milk. Primary and secondary antibodies were diluted in TBST and 5% milk. Blots were scanned using a LI-COR Odyssey Infrared Imaging System, and bands were quantified using ImageJ software. To normalize protein expression levels, densitometric analysis was carried out to housekeeping proteins, beta-actin and tubulin and when required a positive control was also used.

3.17 DNA mismatch repair and CIMP analysis

DNA mismatch repair was analysed through immunohistochemical staining for **DNA mismatch repair proteins** and Microsatellite instability (MSI) molecular tests as already described.[54] Briefly, detection was performed on tumour and when available adjacent normal tissues. The antibodies used (clones, dilutions, antigen retrieval and manufacturer) included: MLH1 (ESO5, 1:100, Bond Epitope Retrieval Solution 2, Novocastra Laboratories, Newcastle, UK); MSH2 (25D12, 1:80, Bond Epitope Retrieval Solution 2, Novocastra Laboratories, Newcastle, UK); MSH6 (44, dilution 1:25, Bond Epitope Retrieval Solution 1, Cell Marque, Rocklin, USA); PMS2 (MRQ-28, 1:20, Bond Epitope Retrieval Solution 2, Cell Marque Rocklin, USA); Tumours were defined as mismatch repair-proficient whereby neoplastic cells showed nuclear immunopositivity for all four

markers. Immunopositive stain was defined as an unequivocal nuclear labelling of the tumour cells with staining intensity comparable to that of internal control. Other internal controls were represented by positive stromal/lymphoid cells. Inflammatory and stromal cells adjacent to neoplastic cells served as positive internal controls. The tumors were defined as mismatch repair-deficient when they showed loss of nuclear "complete absence" staining in at least one of the four markers with concurrent positive labelling in internal non-neoplastic tissue. A recurrent positive expression for MMR proteins was present in normal crypts, fading out towards the lumen. MSI analysis was performed for the National Cancer Institute recommended microsatellite marker panel BAT25, BAT26, D2S123, D5S346, and D17S250 using fluorescently labeled primers on a 3130xl Genetic Analyzer (Life technology, Carlsbad, CA, USA). MSI-H was diagnosed if instability was evident at 2 or more markers as already described.[54] Genomic DNA isolation and sodium bisulphite modification were carried out as reported. The converted DNA was subjected to quantitative methylation specific PCR as reported. [1,30] The following genes (*RUNX3, IGF2, SOCS1, NEUROG1, CDKN2A* (p16) and h*MLH1*) with methylation levels greater than 15% were considered positive. Tumors with at least three methylated loci were classified as CpG island methylator phenotype (CIMP)-positive and the remaining cases as CIMP-negative. The primers for promoter methylation analysis have already been reported.[6,54]

3.18 Interphase fluorescence in situ hybridization (I-FISH)

Fluorescently Labeled Polynucleotide Probes for alpha-satellites of chromosome 17 centromere (CEP17) probe and for the human ERBB2 gene were used as initial screening for detecting chromosomal numerical abnormalities in the in the FFPE tumour sections. The zytolight SPEC ERBB2/CEN17 dual color Probe comprehensive of centromere of chromosome 17 probe (ZyOrange: Orange excitation at 547 nm and emission at 572 nm, similar to rhodamine) and ERBB2 probe (ZyGreen: excitation at 503nm and emission at 528 nm, similar to FITC) were used according to the manufacturer's recommendations (ZytoVision, Bremerhaven, DE). The slides were examined using an Olympus BX61 (Olympus, Hamburg, Germany) with appropriate filters for and for DAPI nuclear counterstain. The alpha-satellite chromosome-specific DNA centromeric probes for chromosomes 1 and 12 centromere (CEN1 and CEN12) were used to further investigate chromosome number into nuclei of the cells (ZytoVision, Bremerhaven, DE). The results were interpreted following enumeration of the signals in 500 interphase nuclei.

3.19 In-house control colorectal carcinoma data sets and Tissue Microarrays

Two independent subsets of patients with primary sporadic colorectal carcinoma (CRC) were included in this study and used as control population. CRC tissues including (n=141, stage I-IV, cohort A) and (n=102, stage IV, cohort B) cancer along with the corresponding adjacent non-involved tissues (n=61) and (n=22) respectively were collected. For both cohorts were available FFPE blocks, while, for 80 out of 141 patients of the (Cohort A) were also available snap-frozen paired tumour-normal tissues in liquid nitrogen and stored at −80°C. Patients who had a familial history of intestinal dysfunction or CRC, had received chemotherapy or radiation before resection or had taken non-steroidal anti-inflammatory drugs on a regular basis were not included. For each patient, the date of colon cancer diagnosis, date of last follow up, and vital status at last follow-up (i.e., living or deceased), Overall Survival (OS) and Progression Free Survival (PFS) were recorded. The latest update was performed on May 1st, 2015. The main clinical-pathologic features of all patients are summarized in (**Tables 6,7).** Colon cancer tissue microarrays TMAs fully annotated with clinical and pathological information were obtained from the two data sets and screened for markers of colon epithelial differentiation (CDX2, CK20, TP53), mismatch repair proteins (MMR)(MLH1, MSH2, MSH6 and PMS2) and genetic *KRAS* and $BRAF^{V600E}$ mutation status. TMAs were assembled as already described.[29,54] Briefly, Cores measuring 0.6 mm in diameter were made in triplicate from paraffin blocks comprising tumor tissues or matched normal mucosa and processed using the ATA-27 automated arrayer (Estigen, Tartu,EW). H&E from TMA were reviewed to ensure that from each case were morphologically similar to those of the corresponding whole tissue section and represented cancerous or normal epithelial cells. All FFPE blocks, frozen tissue and TMA were analyzed.

Table 6. Correlation between CROCC IHC, Clinicopathologic and molecular features of 140 colorectal cancers Stage I-IV.

Parameter		CROCC IHC						P value
		Loss (<1)		Norm (1-2)		Amp (>2)		
		N	%	N	%	N	%	
Gender	Female	3	60	53	63.8	22	41.5	0.435
	Male	2	40	30	36.1	31	58.5	
Localization	Right	3	60	28	33.7	10	18.9	0.0533
	Left	2	40	55	66.2	43	81.1	
Stage	I	0	0	7	8.4	5	9.4	0.211
	II	0	0	35	42.1	25	47.1	
	III	2	40	10	12	10	18.8	
	IV	3	60	31	37.3	13	24.5	
Histology	ADC	2	40	68	81.9	38	71.6	0.0646
	ADC-muc	3	60	15	18	15	28.4	
CK20	Neg	4	80	23	27.8	10	18.9	0.0036**
	Pos	1	20	60	72.2	43	81.1	
CDX2	Neg	4	80	10	12	3	5.6	<0.0001***
	Pos	1	20	73	87.9	50	94.3	
MMR	Neg	2	40	17	20.5	4	7.5	<0.047*
	Pos	3	60	65	79.5	49	92.4	
TP53	Neg	1	20	21	25.3	13	24.5	0.212
	Low	4	80	25	30.1	18	33.9	
	High	0	0	37	44.5	22	41.5	
KRAS	WT	4	80	54	65	33	62.2	0.721
	MUT	1	20	29	34.9	20	37.7	
BRAF (V600E)	WT	2	40	79	95.1	50	94.3	<0.0001***
	MUT	3	60	4	4.8	3	5.6	
Total		5	100	83	100	53	100	

Abbreviations: Right includes: Proximal caecum, ascending and transverse colon. Left includes: Distal descending, sigmoid colon, rectum. ADC adenocarcinoma, ADC-Muc adenocarcinoma with a mucinous component below 50%. Tumors deficient or proficient in DNA mismatch repair (MMR) were identified based on detection of MLH1, MSH2, PMS2 and MSH6 proteins. CROCC expression by IHC was calculated as follows: the number of visible centrosomes was counted on (>200 cells), and expressed as loss (< 1

centrosome per cell); Normal (from 1 to 2 centrosome per cell) or amplified (more than two centrosomes per cell).

Table 7: **102 metastatic Colorectal cancer tissue microarrays investigated by immnohistochemical staining**.

Characteristics	No.	%
No. of Patients	**102**	
AGE Median	64	
SEX		
Male	53	52
Female	49	48
STAGE AT DIAGNOSIS		
II	21	20.5
III	29	28.5
IV	52	51
PRIMARY TUMOR SITE		
Colon	74	72.5
Rectum	28	27.5
SITE OF METASTASES		
Liver	83	81.0
Lung	35	34.3
Others	46	44.9
NUMBER OF METASTATIC SITES		
1	49	48.0
2	44	43.0
>2	9	9.0
K Ras STATUS		
Wild Type	41	40.2
Mutant	50	49.0
Not Available	11	10.8
BRAF (V600E) status		
Wild Type	90	88.0
Mutant	7	7.0
Not available	5	5.0
Mismatch repair (MMR) STATUS		
MMR proficient	83	81.0
MMR deficient	14	14.0
Not available	5	5.0

Abbreviations: others metastatic sites include (Peritoneum, Lymphonodes, Bone, CNS). Tumors deficient or proficient in DNA mismatch repair (MMR) were identified based on detection of MLH1, MSH2, and MSH6 proteins

3.20 Cell culture

Human colon cancer cell lines HCT116, HT29, CaCo-2, LoVo, RKO, T84, DLD1, SW480 and SW620 were purchased from American Type Culture Collection (ATCC, Rockville, MD, USA) and cultured as appropriate at 37 °C in 5 % CO2 atmosphere in Dulbecco's modified Eagle's medium, DMEM (Thermo fisher scientific, Whaltam, MA, USA) or RPMI 1640, (Thermo fisher scientific, Whaltam, MA, USA) supplemented with 10 % fetal bovine serum (FBS) or when required without FBS, 100 U/ml penicillin and 100 ng/ml streptomycin (Thermo fisher scientific, Whaltam, MA, USA). BJ human skin fibroblasts derived from normal foreskin and G401 cells derived from pediatric rhabdoid tumor were maintained in DMEM and RPMI 1640 medium, respectively, supplemented with 10 % fetal bovine serum (FBS), 1 mg/mL penicillin, 1 mg/mL streptomycin in a humidified incubator at 37 °C and 5% CO2. All cell lines were tested with the mycoAlert Mycoplasma Detection Kit (Lonza Walkersville, Walkersville, MD, USA) and found to be free of mycoplasma. The genotype of parental cell lines were confirmed by short tandem repeat (STR) genotyping. The STR genotypes were consistent with published genotypes for each cell lines. All cell lines were found to be free of mycoplasma tested with the mycoAlert Mycoplasma Detection Kit (Lonza, Basel, Switzerland).

3.21 Plasmid transfection

The RKO cells at 70-80 % confluence was transiently transfected with SureSilencing control or CROCC shRNA expression plasmids (KH23140P, (Qiagen, Hilden, DE) containing the puromycin resistance cassette. Transfected cells were selected with 0.8 μg/ml puromycin (Thermo fisher scientific, Whaltam, MA, USA) for 1 week and single colonies were amplified and assessed for efficient CROCC silencing by semiquantitative and quantitative PCR (qPCR). For rescue experiment, the full-length CROCC coding sequence "clone 6150861 pEGFP Rootletin, Nigg pFL2(CW499)" gift from Erich Nigg or a truncate form (1–494aa) cloned with GFP epitope or GFP alone (used as control) were transfected in metastatic T84 cells and selected in 2 mg/ml G418 before flow-sorting CROCC-GFP expression.[58] Cell vitality for long-term experiments CROCC-GFP+ cells were maintained in 0.6 mg/ml G418. All transfections were performed with

Lipofectamine 3000 (Thermo fisher scientific, Whaltam, MA, USA) according to the manufacturers' recommendations.

3.22 Metaphase spreads and clonal FISH preparation

For metaphase spreads, CRC cells were collected after 1-h treatment with 10mM colcemid Gibco® KaryoMAX® Colcemid™ (Life technology, Carlsbad, CA, USA), (Gibco), and swelled with KCl (0.4%, 37 °C, 7 min) before fixation in 3:1 methanol:acetic acid. Cells were dropped onto glass slides and aged for,2 weeks. For clonal FISH, 500 cells were expanded into colonies of 30–60 cells on glass slides before KCl treatment and fixation in 3:1 methanol:acetic acid. Slides were denatured (70 uC in 23 sodium citrate (SSC)/75% formamide, 2 min, quenched in ice-cold 70% ethanol) and dehydrated through an ethanol series. Chromosome specific Alpha-satellite repetitive centromere DNA for chromosome 1 and X, respectively were used (CEP1, CEPX Abbott, Abbott Park, IL, USA). The probes were denatured (90 uC, 6 min) and hybridized to slides (16 h, 37°C), and then washed. Slides were dehydrated and mounted in Vectashield hardset plus DAPI mounting medium (H-1500).

3.23 Nocodazole washout assay

Cells were then incubated for 5, 15 and 30 min with the microtubule destabilizer nocodazole (10 μg/ml) at 37°C, washed five times with PBS at room temperature. To determine recovery times, slides were then fixed in -20°C methanol. Microtubule structures were detected using antibodies described above against anti–gamma-tubulin (1:500 in 1% BSA for 1 hr) and/or CROCC (1:200 in 1% BSA for 1 hr). The primary antibody was detected with an FITC labeled sheep antimouse antibody (1:500, 1 hr, Sigma-Aldrich, Saint-Louis, MO, USA) and cells counterstained with propidium iodide. Asters were counted when clear signals at the periphery of the nucleus were detected. The completion of microtubule regrowth was scored when the pattern of immunofluorescence resembled the one prior to the nocodazole induced depolymerization. The amount of time required for the reformation of microtubules from the endogenous tubulin was assessed for normal human fibroblasts, BJ cells used as control.

3.24 Immunofluorescence on CRC cell lines

Cells grown on coverslips were rinsed in PBS and fixed with 4% paraformaldehyde at room temperature for 20 min. After rinsing with PBS, they were incubated with 0.1% Triton X-100 at room temperature for 10 min to permeabilize the cell membrane and rinsed again with PBS. Cells were incubated in blocking solution (6% BSA in PBS) for 1 h at room temperature and then with

primary antibody overnight at 48C. The primary antibodies were rabbit polyclonal or mouse monoclonal anti-γ-tubulin,α-tubulin and CROCC as for immunohistochemistry. The slides were rinsed three times with PBS and incubated for 1 h at room temperature with secondary fluorochrome-conjugated antibodies anti-rabbit AlexaFluor 594 and anti-mouse AlexaFluor 488, respectively. DAPI was used to label cell nuclei. The preparations were mounted with Vectashield mounting medium (Vector Laboratories, Burlingame, CA, USA) and examined with an Olympus IX81 deconvolution fluorescence microscope (Olympus Microscopes, Center Valley, PA). No staining was detected when conjugates alone were used as negative controls. Alternatively, cells grown on coverlips were fixed in cold methanol and double-label stained with monoclonal anti-γ-tubulin and anti-CROCC or in addition anti-γ-tubulin anti-α-tubulin antibodies. A secondary anti-mouse antibody, conjugated with Cy3, and an anti-rabbit antibody, conjugated with FITC, were diluted 1:500, 1:1000 respectively.

3.25 Immunoblotting

Briefly, protein extracts from cell lines and tissue sections of fresh tissue specimens from tumour and matched normal adjacent mucosa frozen in liquid nitrogen were prepared and analyzed as previously reported.[54] The primary antibodies included those described above; CROCC, gamma-tubulin, beta-tubulin, β-actin and SMARCB1 (H300, Santa Cruz, Dallas TX, USA). The secondary antibodies were used anti-mouse (sc-2031) and anti rabbit (sc-2004) (Santa Cruz, Dallas TX, USA).

3.26 CAsy cell counter and Proliferation assay

Cell number and density of viable cells were determined using CAsy Cell Counter (Roche Innovatis AG, Reutlingen, DE). Each sample (cell suspension) was prepared three times in CAsyTon buffer (Roche Innovatis AG, Reutlingen, DE), followed by triplicate measurements of 200 µl sample volume. All counts of a size smaller than 10 µm (dead cells and debris) were excluded. For cell proliferation assay, 5,000 cells/well were seeded in 96-well plates. The number of proliferating cells was evaluated by 3-(4,5-dimethyl-2- thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT). At the indicated times, MTT solution (Sigma-Aldrich, Saint-Louis, MO, USA) in complete medium (0.28-mg/ml final concentration) was added and incubated at 37°C for 4 h. The medium was discarded, and the formazan salts were dissolved in 4 mM HCl, 0.1% NP40 in isopropanol. The colorimetric substrate was measured and quantified at 560 nm in an enzyme-linked immunosorbent assay plate reader.

3.27 Cell morphology, wound healing and invasion assay

Initially, cell motility was evaluated by the wound-healing assay. Briefly, cells were grown to confluence and a wound made through the monolayer using a p1000 tip. Accurate measures of the wounds were taken during the time course to calculate the migration rate according to the equation: *percentage wound healing = ((wound length at 0 h) - (wound length at 24, 48 or 72 h))/(wound length at 0 h) x 100.* Phase-contrast images were acquired every 2 h for 24 h and the percentage of migrating cells was determined automatically. Three independent experiments were performed. For invasion assay, 2×10^4 cells were added to the upper compartment of a 24-well BioCoat Matrigel Invasion Chamber (Corning Incorporated, Tewksbury, MA, USA) in serum-free DMEM. After 24 h, invading cells were fixed, stained with crystal violet 0.1% and counted. Hematoxylin&Eosin staining was used to reveal morphological changes in CRC cells under investigation.

3.28 Cell cycle and Flow cytometry analysis

Cell cycle analysis was performed three days after seeding on both attached and floating cells using the BD Cycletest Plus DNA reagent Kit (Cat. No. 340242; (BD Biosciences San Jose, CA, USA). Propidium Iodide stained cells (>20.000 events) were analyzed by flow cytometry on FACSVerse (BD Biosciences San Jose, CA, USA). Debris and doublet cells were excluded and only single cells were considered for cell cycle analysis. Results were reported as percentage of cells in G1, S and G2/M phases. All flow cytometry results were analyzed with the FACSuite Software v.1.0.5.3841 (BD Biosciences San Jose, CA, USA). Four biological replicates were prepared and each assayed in triplicate; the results were reported as mean ± standard deviation (SD).

3.29 Statistics

Patient subsets were compared for survival outcomes, using both Kaplan-Meier survival curves and multivariate analyses based on the Cox proportional hazards method. Differences in Kaplan-Meier curves were tested for statistical significance using the log-rank test even once stratified based on gene or protein expression patterns. For meta-analysis on the rhabdoid colorectal cancer, data from 19 reports in the literature were compiled. This resulted identification of a total of 23 patients for whom follow-up and some molecular data were available for 20 and 9 patients, respectively **(Table 8).** Life table analysis was performed in order to provide Kaplan-Meier estimates of survival after stratification separately for clinicopathological factors. Differences in survival among subgroups were analyzed using a Cox proportional hazards analysis on other variables. Relative risks of death, ratio of mortality rates between subgroups, were estimated from the Cox regression model with and without adjustment for other factors. Progression Free Survival (PFS) was compared between the groups when available and defined as the time elapsed between the start of first line chemotherapy

and disease progression, treatment discontinuation or death. Overall Survival (OS) was defined as the time elapsed between the start of first line chemotherapy and death. The pattern of gene expression profile *(GEP)* in independent data sets "TCGA pooled series" of patients was analyzed as previously reported.[54,59] The presence of an enrichment defined tumors subtypes characterized by high pathological grade (G3/G4), microsatellite, instability (MSI) or gene mutations was tested using Pearson's t-test and by computing odds-ratios (OR) together with their 95% confidence intervals (CI). Differences in the expression levels of individual genes or other markers among different sample subgroups were evaluated using boxplots and tested for statistical significance using a 2-sample t-test (2-tailed). Patients with documented follow-up were also available for survival analysis. We censored those patients who were alive without tumor recurrence or dead at last contact. Wilcoxon-Mann-Whitney and Kruskal-Wallis tests with median differences at 95% confidence interval (CI). The Spearman rank test was used to assess the correlation between continuous variables, and the Pearson X^2 test for the association between categorical variables. Data are presented with mean, medians and ranges. The *P* values were calculated two sided. Statistical analyses were conducted were performed by GeneSpring R/bioconductor v.12.5 and R based package, SPSS (version 15 Windows, SPSS Inc, Chicago, IL), and GraphPad Prism 5.

Table 8. Rhabdoid colorectal cancer reported in literature between 1993-2015

Number	Case	Age	Sex	Location	metastasis	type	survival (m)	(BRAF, KRAS and MMR)	SMARCB1 (IHC)
1	Baba	45	F	?	?	?	1.5	ND	ND
2	Romera	77	M	DC	no	pure	2	ND	ND
3	chetty	72	F	Cecum	yes	composite	?	ND	ND
4	Yang	75	M	Colon	no	pure	0.5	ND	ND
5	markus	84	F	Colon	No	composite	? alive	ND	ND
6	Nakamura	76	M	Cecum	yes	pure	2	ND	ND
7	Kono	66	M	Cecum	no	composite	1.5	ND	ND
8	Mastoraki	62	F	Colon	yes	pure	4	ND	ND
9	Seok	63	M	Cecum	yes	pure	?	ND	ND
10	Hoon-kuy	69	F	Sigmoid	no	composite	6	ND	ND
11	Lee	62	M	Sigmoid	no	composite	36 alive	ND	ND
12	Lee	83	M	Rectum	yes	composite	1	ND	ND
13	Agaimy	79	M	Cecum	no	composite	6	Yes (BRAF/D)	Neg
14	kalyan	31	F	Cecum	yes	composite	4	Yes (KRAS/P)	++
15	Remo	73	F	Cecum	no	composite	6	Yes (BRAF/D)	Neg
16	Pancione	71	F	Cecum	yes	pure	8	Yes (BRAF/D)	++
17	Stulpinas	49	M	Rectum	yes	composite	7	Yes (BRAF/P)	++
18	Moussaly	87	F	Transverse	no	composite	2	ND	ND
19	Cho	73	M	Cecum	no	composite	1 alive	ND	ND
20	Sanchez	77	M	Transverse	no	pure	2	Yes (WT/P)	ND
21	Sanchez	65	M	Descending	no	pure	12 alive	Yes (WT/P)	ND
22	Sanchez	63	M	Descending	no	pure	1	Yes (KRAS/P)	++
23	Sanchez	71	F	Cecum	no	composite	8	Yes (BRAF/D)	++

All RC reported are characterized by loss of CK20 and positivity for Vimentin (IHC).

Molecular test BRAF (V600E), KRAS, MMR: ND: Not Done (14/23=60%);

Location: cecum: (10/23=43%)

Type: composite; (13/23= 56%)

MMR= D, Defective; P, Proficient

4 Results

4.1 Identification of *CROCC* mutations in two rare rhabdoid colorectal cancers.

To decipher the molecular pathogenesis of these disease entities, two rare cases of primary RCs (RCI and RCII) harboring MSI due to *MLH1* promoter methylation, CIMP, *BRAF* V600E mutation and wild type *SMARCB1* [6,19] were subjected to whole exome sequencing (WES) using DNA from formalin-fixed-paraffin embedded matched tumor/normal samples. Consistent with MMR defects but not with genomic profiling of the ERTs, we detected an exceptionally large number of somatic point mutations 1056 and 1078 per 10^6 bases for RCI and RCII, respectively **(Figure 1)**. [41,60,61]

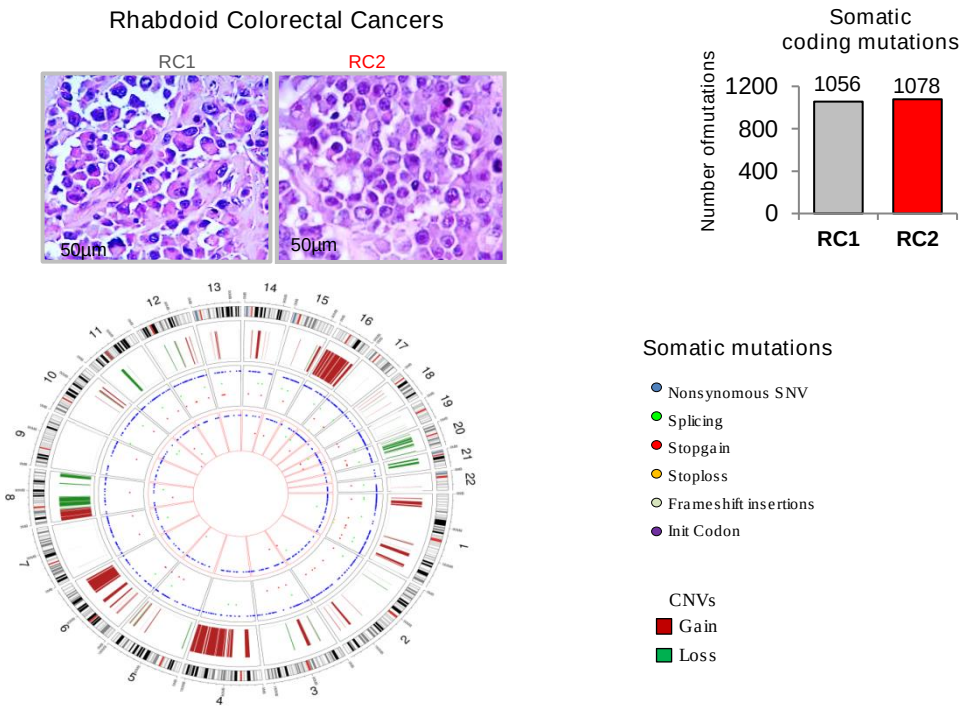


Figure 1. Whole exome sequencing reveals CROCC mutations in two patients with rhabdoid colorectal cancer.
Representative H&E pictures of two rare rhabdoid colorectal cancer specimens (RCI and RCII patients) sequenced by whole exome sequencing. The upper graph indicates the total number of somatic mutations per tumor. The distribution of non-silent single nucleo tide variants (blue), and other common non- silent mutations splicing (green) stop gain (red) are shown using circos. The outer rings are chromosomes, the next ring depicts copy number (red represents gain and g reen represents loss)

The distribution of somatic substitutions showed a higher rate of transitions as compared to transversions (71.8% vs 28.2%) with a dominance of C>T/G>A, T>C/A>G transitions. About 1/5 of mutations occurred within CpG dinucleotide context as has been seen in other classical colorectal cancers (CRCs).[41,61] The relatively elevated frequency of C>G/G>C and C>A/G>T transversions also supported a mutational signature involving DNA repair mechanisms other than h*MLH1* deficiency**(data not show)**. However, as has been found in most MMR defective tumor types, the most prevalent single nucleotide variants (SNVs) were non-silent mutations[41], whereas, over 90% of potentially damaging mutations were missense and around 10% were splicing, stop-gain, stop-loss or rarely frameshift insertions or initiation codon mutations (**Figure2A)**. We identified 112 (10%) shared protein-coding mutations which enriched biological processes of relevance to the hallmark phenotypes (**Figure 2B)**.

A

GENETIC ALTERATION	RCI	RCII
missense	961	1003
splicing	33	22
stop gain /Fs ins/start cod	47	39
stop loss	4	0
unknown	11	14
Total	1056	1078

B

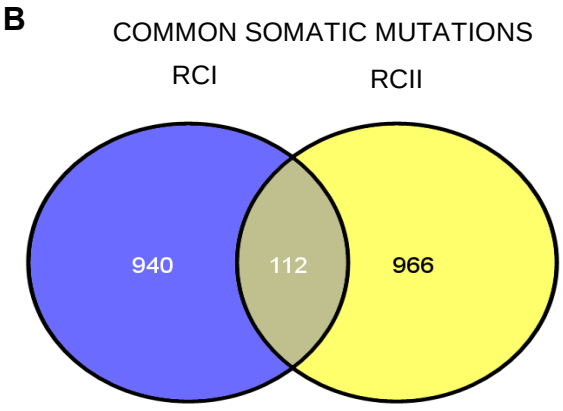


Figure 2 mutational analysis. A)Number and type of mutations reported in RC1 and RC2. B) shared protein-coding mutations

By applying a bioinformatics tool, "DrGaP", [23] to the shared somatic mutations, we found 20 potential candidate disease-causing genes, of which nearly half (45%) were related to cytoskeleton/centrosome and microtubule biological functions **(Figure 3)**

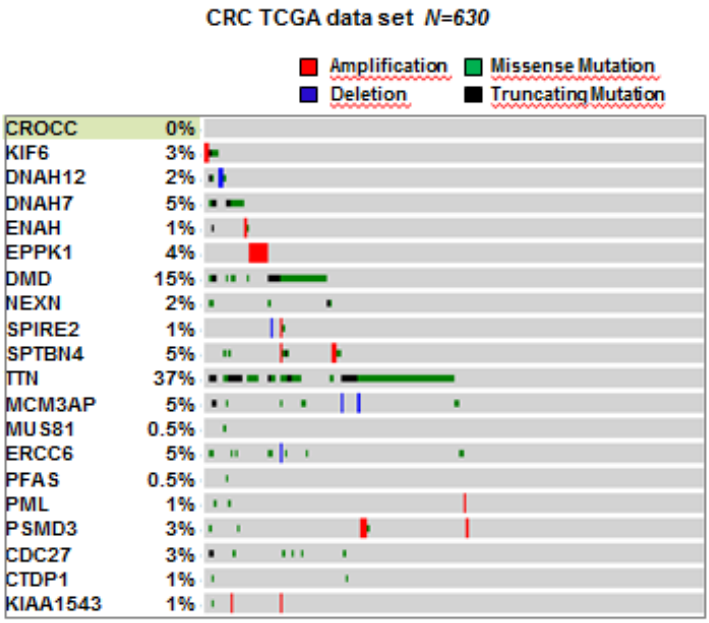
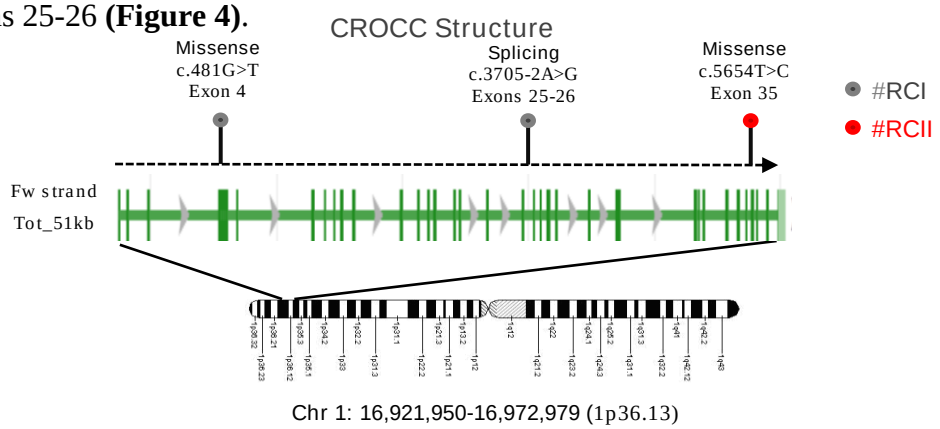


Figure 3. Recurrence of twenty shared potential diseases-causing genes in (n=630) colorectal cancer TCGA database

.

The search in public TCGA database[41,61,62] comprising 630 classical CRC (http://www.cbioportal.org), revealed that the majority (13 out of 20; 65%) of the 20 candidate genes had low frequency of mutations (≤4% of cases) and strikingly among these, only one candidate *CROCC* "rootletin" [58], a centrosome linker gene mapping to 1p36.13 was notable, because no somatic mutations (0/630; 0%) were reported**.** We discovered two missense mutations in *CROCC*, p.Ala161Ser (c.481G>T, Exon 4) and p.Val1885Ala (c.5654T>C, Exon 35), and one remarkable splicing mutation (c.3705-2A>G) at the conserved 3' acceptor splice site in the intron between exons 25-26 **(Figure 4)**.



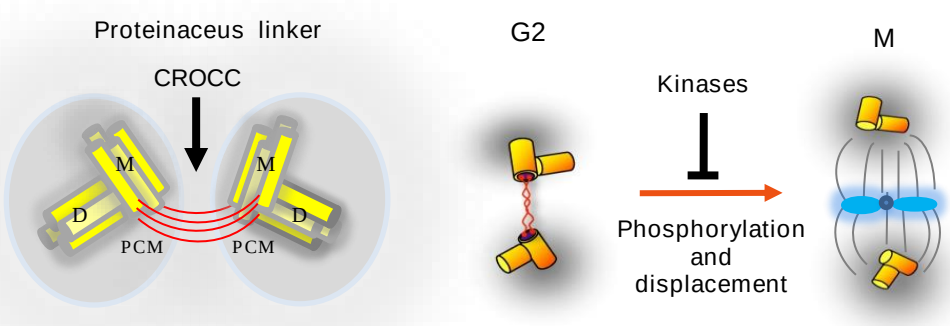


Fig4. chromosome localization (1p36.13) and organization of CROCC or "rootletin"
All exons 37, are depicted as green vertical bars and introns as horizontal lines. Solid circles indicate mutations identified in two RCs. CROCC forms the structural basis of centriole-centriole cohesion system "proteinaceous linker" black arrowhead. Centrosome consists of two pairs of centrioles daughter centriole and mother centriole surrounded by the pericentriolar material (PCM). Before mitotic entry, the fibrous "proteinaceous linker" is degraded through mitotic kinases to allow centrosome separation and define the opposite spindle poles for mitotic spindle assembly (M)

Notably, the splicing mutation verified by sanger sequencing in patient (RCI), reduced the strength of the physiologic acceptor site, predicting a large deletion of the *CROCC* coding region (-32.51% of the wild-type) involving exons 23-31 (**data not show).** Next, RT-PCR amplifications "spanning *CROCC*-exons 5–7 and 33-35" were carried out using cDNA from the paired tumor-normal mucosa. The tumor bearing spicing mutation**,** exhibited low CROCC mRNA and protein expression which prevented us from building a exon junction map, then indicating alteration of the mature transcript by the utilization of cryptic splice sites or by the activation of the nonsense-mediated mRNA decay pathway, which impair transcripts harboring large deletions[63] **(Figure 5).**

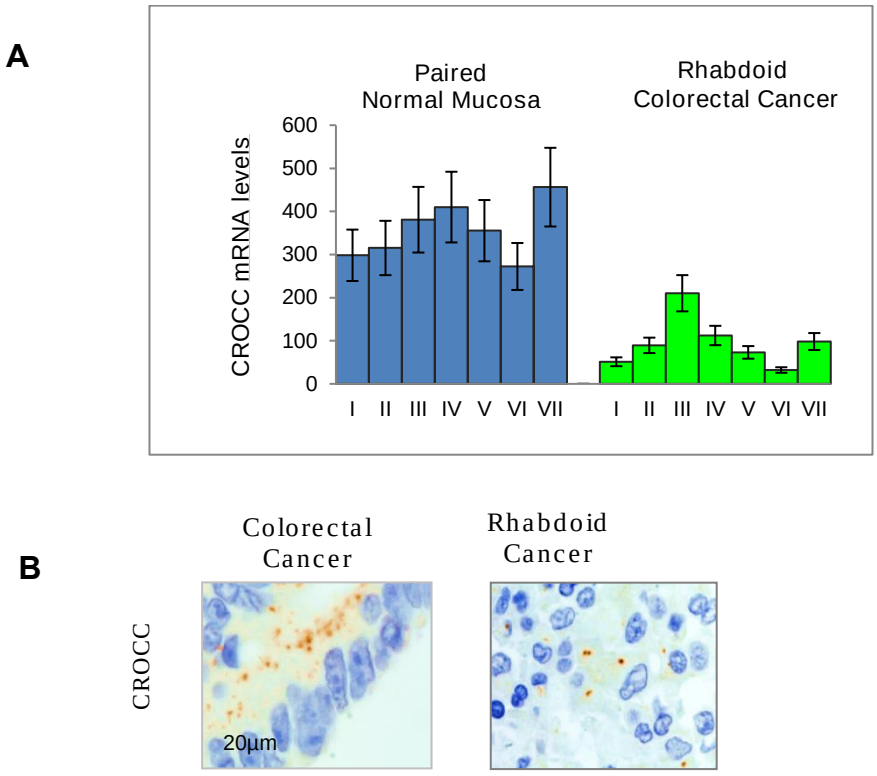


Fig5 Patient-matched tumor-normal expression quantification of CROCC mRNA (A) identifies consistent, downexpression in all seven RCs. *P < 0.01 derived from two-tailed Student's t test for each patient. (B) In normal colon epithelia CROCC labeling reveals round and uniform in size centrosomes, whereas in rhabdoid cells centrosomes are reduced in number or display profound alterations resulting in mispositioned or larger centrosomes into anucleated cells, "inset modeled image"

Patient-matched tumor-normal expression from individual RCII, harboring the missense mutation (c.5654T>C) also confirmed the previous observation with lower *CROCC* expression levels in tumor than in normal tissue (data not show). **T**he screening across major cancer types revealed that *de novo* mutations in the *CROCC* gene were poorly represented across NCBI/TCGA database, whereas, the majority of tumors including pediatric rhabdoid derived cell lines[64] displayed recurrent overexpression of *CROCC* (45%, 9 out of 20) and a few of genetic deletion at 1p36.13 locus (10%, 2 out of 20) **.**[61] By contrast, we found recurrent copy number loss at *1p36.13* locus[65] in neuroblastoma (65%, 11 out of 17) and in a small but relevant proportion of aneuploid "MSS" (23.6%, 9 of 38) colon cancer cells than in non-aneuploid "MSI" (4.5%, 1 out of 22**.** Compared to CRC cells retaining *1p36.13*, those harboring the deletion revealed a gene-expression signature significantly enriched for pathways implicated in chromosomal instability[66,67,68](**data not show).** Thus, we suspected that deleterious mutations affecting *CROCC* at the centrosome underlie RC pathogenesis in two patients.

4.2 Consequences of *CROCC* dysfunction in rhabdoid tumors.

In an effort to analyze the significance of the mutations identified in the discovery screen on the two cases, we collected further five patients with RC, thus a total of 7 cases (4 males and 3 women, mean age 67 years) were studied **(Table 1).** Notably, across the five RCs, we did not identify additional mutations in *CROCC* but loss of heterozygosity (LOH) at the *1p36.13* locus associated to mRNA reduction, below normal levels in all cases **(Fig. 6 and Table 1).**

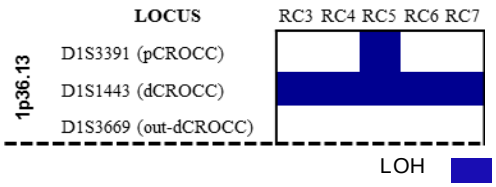


Fig 6. loss of heterozygosity analysis (LOH) at the *1p36.13* locus in rhabdoid colorectal cancer

These results suggested recurrent lack of *CROCC* expression as a consequence of the loss-of-function in the wild-type allele[65]. The majority of the RCs revealed *BRAF*^{V600E} mutation, CIMP and

a wild-type *SMARCB1*, while, "3 out of 7 tumors" had MSI associated with promoter hypermethylation of *MLH1*, and the remaining 4 cases were MSS or mismatch repair–proficient tumors **(Table 1)**. We next aimed to determine the level of mitotic aberrations and found that all RCs displayed a high prevalence of bizarre mitotic figures, prevalently segmented/multi-lobular nuclei or aberrant enucleated cells and a variable degree of cytomorphologic and ploidy aberrations such as triploid or near-tetraploid cells ranging from 10% to 40% of tumor cells **(Figure7)**.

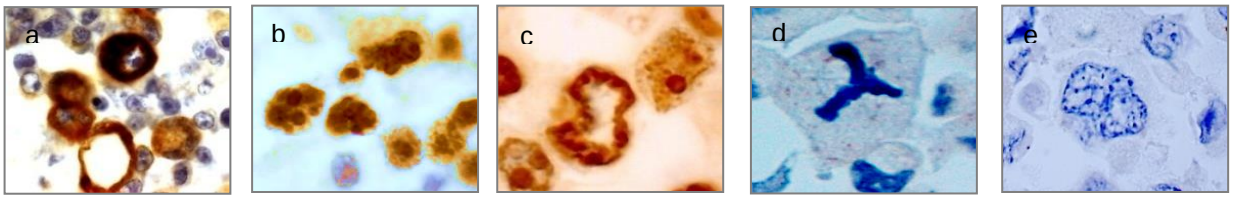


Figure7. from the left, representative microphotographs including anucleated cells, cytomorphologic aberrations(a) including multi lobulated nuclei(b), abnormal mitotic figures (c), tripolar mitosis(d), mitotic catastrophe with fragmented chromatin(e)

Compared to our dataset of RCs, genetic profile of seven ERTs hereafter named (RI), revealed predominant missense or truncating mutations in *SMARCB1* (5/7, 71%), [15] or in *TP53 (3/7*, 42%) a near diploid DNA content[18] and less aggressive clinical course .[6-11] (**table 9**)

Table 9. Infants/young adults rhabdoid cancers of the central nervous system: clinicopathological and immunohistochemical features of the 7 cases.

Case	Age (years)	Sex	Location	INI1 (IHC)	CROCC (IHC)	Follow up (months)	Status
RI 1	2	F	CNS	-	-	n.a.	n.a.
RI 2	2 months	M	CNS	-	Single, large centriol	3	Dead
RI 3	4	M	CNS	+	Single, large centriol	33	Alive
RI 4	5	M	CNS	-	Single, large centriol	21	Dead
RI 5	10	F	CNS	+	Single, large centriol	92	Alive
RI 6	6	M	CNS	-	Single, large centriol	n.a.	n.a.
RI 7	19	F	CNS	-	Single, large centriol	46	Alive

CNS, central nervous system; AT/RT, Atypical teratoid rhabdoid tumor
n.a., Not Available

We therefore conclude that RCs although morphologically indistinguishable from their pediatric counterparts demonstrate distinct molecular and cytogenetic aberrations. We next determined the consequences of the genetic alterations on the CROCC protein using paired non-tumor tissues as control. Immunohistochemical analysis on RCs, revealed tumor cells with no CROCC labeling (nearly 50%), cells with a single and often incorrectly positioned and/or fragmented positive centrosome with most having bizarre phenotypic defects, for example, enucleated cells harboring a large and single centrosome in the late telophase **(Figure 5b**). The study of γ-tubulin labeling as reference marker to centrosome[58,69] revealed either a diffuse accumulation into cytoplasmic fraction[52] often accompanied by a single centrosome or lack of centrosome staining (~50% of cells) (**Figure 8**).

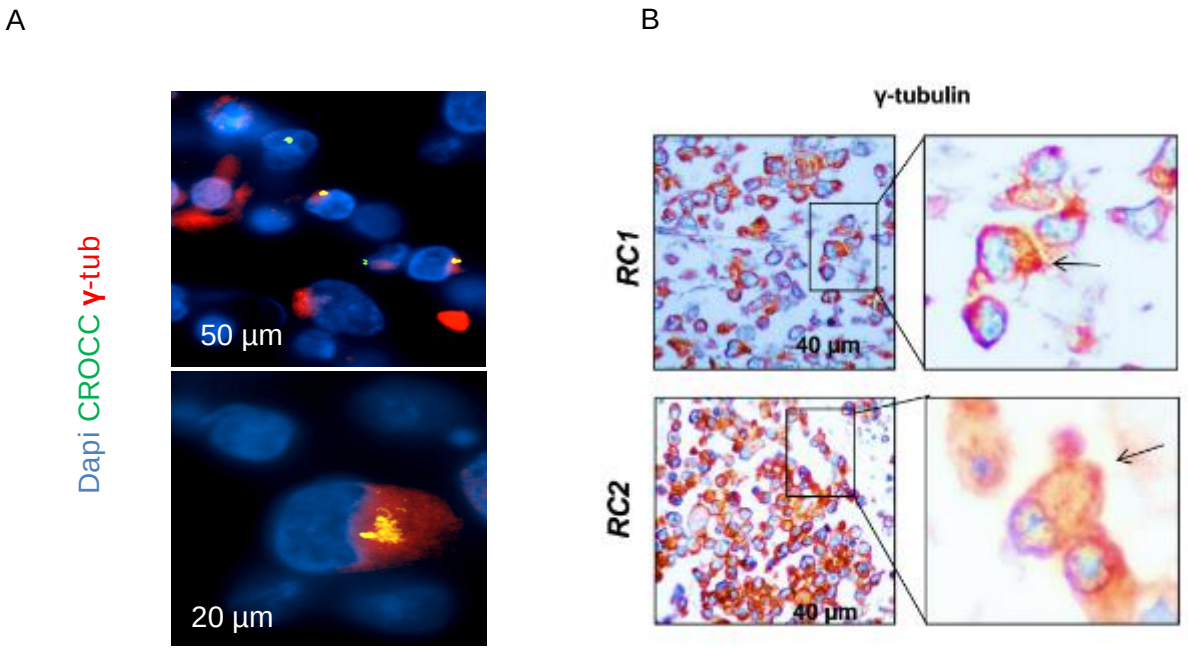


Figure8: γ-tubulin labeling of rhabdoid colorectal cancer tissue a) immunofluorescence merge of crocc (green) and γ-tubulin (red) show a single big centrosome. b) immunohistochemistry of γ-tubulin on two cases of rhabdoid colorectal cancer show mitosis monopolar with single centrosome

In contrast, the large majority of RI predominantly had a single or pair CROCC-positive centrosome (80%) often closer to the nucleus than RCs, but considerably larger, and only (10%) were cells with no centrosome. Compared to RCs, analysis of labeled centrosomes was associated with much higher tumor/normal paired expression of *CROCC* (*P*<0.00001) **(Figure.9).**

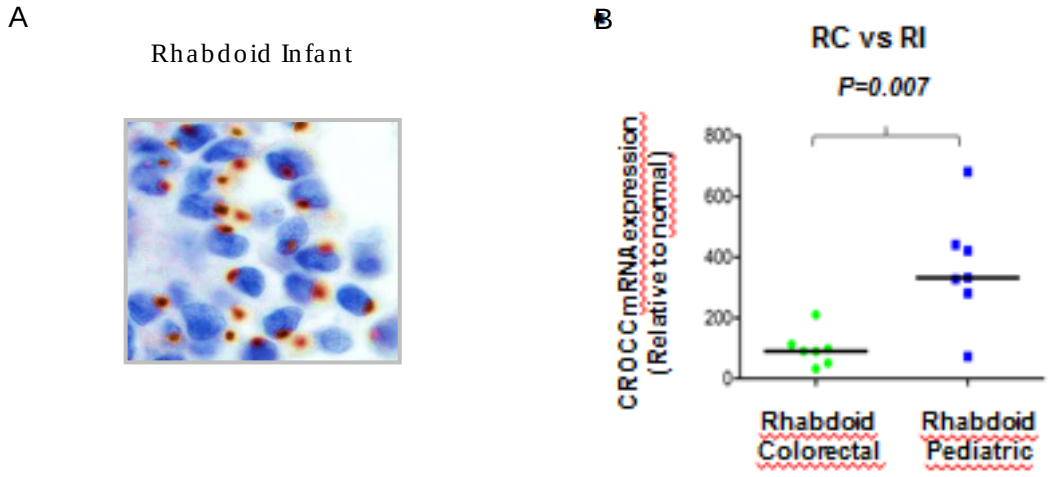


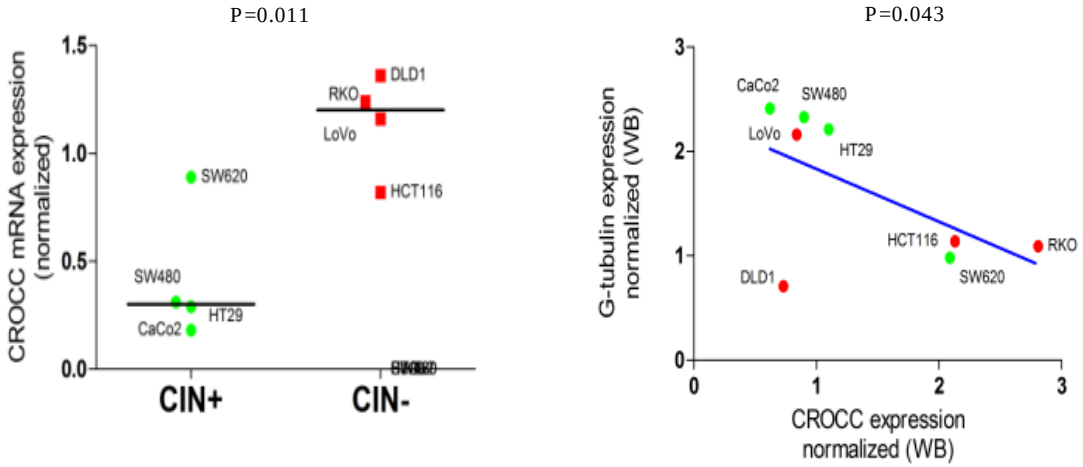
Figure9 A) Larger centrosomes (~1 per cell) close to the nucleus marked by CROCC, are characteristic of pediatric rhabdoid tumors. **B**)Matched tumor-normal CROCC mRNA expression level shows significant differences between rhabdoid colorectal cancer (RC) and pediatric cases (RI), P value was obtained by Mann–Whitney test.

Consistently, we found that *CROCC* expression levels tended to be either unchanged or significantly higher also in a substantial proportion of 140 classical CRCs than in 62 normal matched normal mucosa. Immunohistochemical analysis from the same data set, revealed predominantly centrosomes either in normal number (1/2 per cell, 58.5%) but larger in diameter or often supernumerary[1,3] (>2 per cell, 38%) than those in normal epithelium, while only (<1 per cell, 3.5%) were tumors lacking centrosomes staining. This latter subgroup compared to tumors with normal or supernumerary centrosomes, was not associated with overall survival HR=1.28; 95% CI (0.265-6.23); *P= 0.754*. We also detected a higher percentage of CRCs with supernumerary (50%) or defective (12%) compared to normal centrosomes (38%; *P=0.001*) in an independent data set comprising 102 patients with metastatic colorectal cancer. In this dataset, tumors harboring CROCC deficiency were associated with poorer clinical course than those expressing normal or

supernumerary centrosomes (HR=0.30; 95% CI (0.21-0.81); *P<0.0001*. Overexpression of *CROCC* in CRC was independently validated by querying publically available TCGA mRNA expression data. Although *CROCC* expression was not related with tumor stage and overall survival, we found consistently higher levels in MSI than MSS tumors. However, extended analysis from TCGA database also revealed a few of genomic structural variations (15/616; 2.43%) with a preference for copy number gain at the 1p36.13 locus. Thus, *CROCC* deficiency caused by *de novo* somatic mutations or copy number loss are recurrent in RCs but not in related cancers lacking *CROCC* alterations.

4.3 Functional *CROCC* depletion impairs mitosis and induces rhabdoid phenotype.

In line with findings in CRC dataset, *CROCC* mRNA and protein expression levels, were concordant and significantly higher in MSI cells than in MSS "aneuploid" cell lines *(P*<0.05, **Fig. 10**). Notably, all microsatellite unstable cells contained centrosomes stained for CROCC and γ-tubulin that were functionally and structurally indistinguishable from those in normal human fibroblasts BJ.[70]



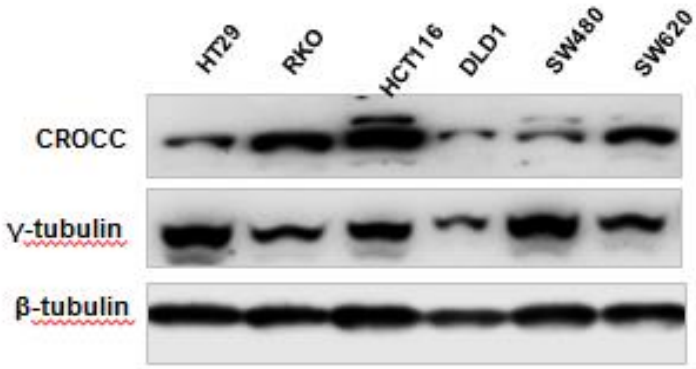


Fig10) Quantitative RT-PCR amplification of CROCC mRNA into MSS versus MSI subgroup colorectal cancer cell lines. RT-PCR analysis was normalized to rRNA 18S. P value was derived from two-tailed Student's t test. Inverse correlation between CROCC and γ -tubulin quantification related to ù-tubulin by western-blot analysis in a panel of eight cancer cell lines classified as MSS (red) versus MSI (green). The P value was derived from Spearman r correlation. Low, immuno-blot analysis showing CROCC and Û-tubulin together in representative cancer cell lines

Although MSS cancer cells, have an increased frequency of micronuclei, prometaphase DNA damage "nuclear γH2AX foci" compared to MSI, we did not detect in any of these cell lines chromosome segregation errors resulting in gross mitotic defect[67,70] . Therefore, we reasoned that RKO cells "MSI" having a near-diploid karyotype and shared gatekeeper molecular alterations (*BRAF* ^{V600E} mutation, CIMP) with RC patients, could be an useful system to explore compromised *CROCC* function *in vitro*. We found that the clone sh4, hereafter named (CROCC^{KD}) provided a stable and consistent knockdown of *CROCC* transcript to more than 75% and protein to 3.3-fold lower then RKO cells transfected with control vector (shCon) achieving nearly comparable levels to those seen *in vivo* **(Figure 11**).

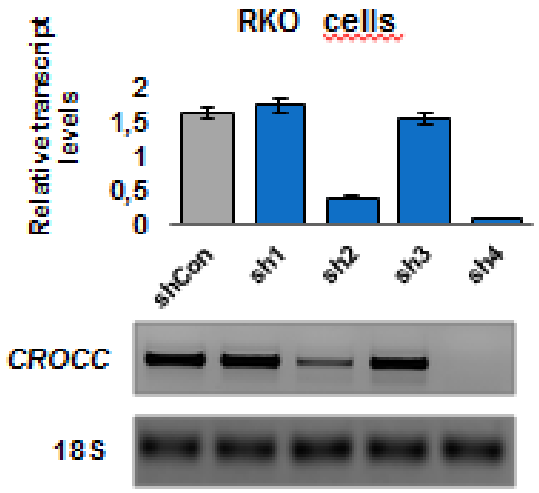


Figure 11. CROCC silencing was performed by transfecting into RKO a pool of short-hairpin (shRNA) plasmids (sh1-4) targeting different region of the CROCC gene or a control shRNA plasmid for nonspecific off-target effect (shCon). Selected cells were assessed for CROCC silencing by semiquantitative RT-PCRand the most efficient (sh4 clone) by western blot analysis.

CROCC depletion caused aberrant mitotic divisions resulting in a higher frequency of monopolar spindles and incorrectly aligned chromosomes as compared to control**.** During metaphase, we observed prominent monopolar spindles with a single large centrosome or with structurally compromised "fragmented" centrosomes, which accounted for 85% of the abnormal phenotype (**Figure 12**).

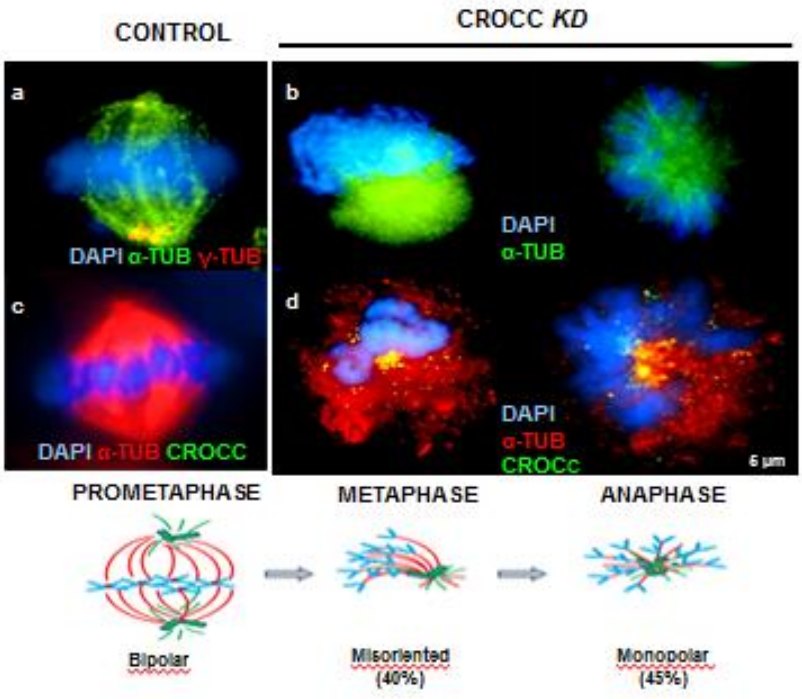


Fig 12) CROCC is stably silenced in mismatch repair deficient RKO cells using short interfering RNA knockdown (CROCC KD). shControl (Control) denotes a non-targeting siRNA control. siRNA-mediated CROCC depletion leads to abnormal spindle formation "misoriented or large monopolar spindles" as compared to control (left panel of images). During mitosis, the monopolar spindles can be subdivided into spindles that had one large or fragmented centrosome. (see supplementary figures). Cells stained for microtubules (α-tubulin), centrioles (γ-tubulin, CROCC) and DNA (dapi) are indicated. Schematic of a bipolar wild-type spindle and the breakdown of the two abnormal spindle phenotypes seen upon CROCC depletion accounting for 85% of the abnormal phenotype.

Consistently, an increased frequency of micronuclei (median 11% $CROCC^{KD}$ versus 1% ShCon cells *P=0.0003*) and γH2AX nuclear foci (median 43 % $CROCC^{KD}$ versus 18% ShCon cells *P=0.011*, was observed. Metaphase karyotyping revealed that *CROCC* deficiency leads to an increased number of tetraploid (4N) cells (median 13.3% $CROCC^{KD}$ versus 3.51% ShCon cells *P=0.001*) characterized by prominent and larger nuclei than diploid cells (2N). Consistently, analysis of centromeric probes in intephase nuclei confirmed tetraploidy. **(Figure 13)**

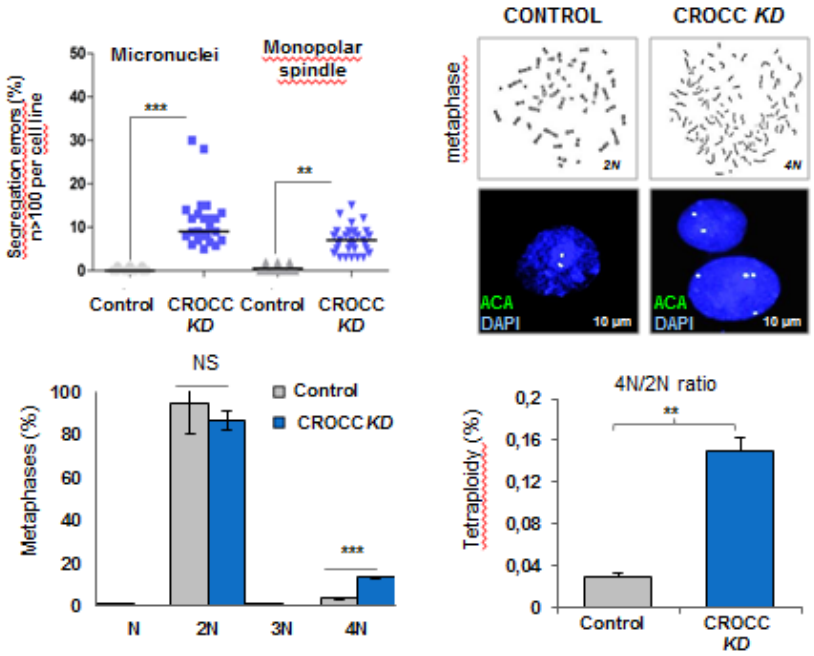


Figure 13. ***CROCC*** **depletion induces rhabdoid phenotype exacerbating mitotic splindle DNA segregation errors through tetraploidy in colorectal cancer** Percentage of DNA segregation errors accounted for by micronuclei and monopolar spindles. P values were obtained by Mann–Whitney test. Representative images of metaphase chromosome spreads of shCon and CROCC KD. Right panels denote example images in cells stained with anti-centromere antibodies (ACAs) and with the 49,6-diamidino-2-phenylindole (DAPI). and Scale bar, 10 µm. On the left, quantification of chromosome number by karyotyping cells (26 chromosome spreads were quantified in each condition). On the right, tetraploid on diploid cells ratio is shown. Error bars represent mean ± s.e. The ***P<0.001 and **P<0.01 were derived from two-tailed Student's t-test.

Moreover, the number of *CROCC*-deficient cells was extremely reduced in G0/G1 or G2/M phases when compared to the wild-type population (by 26–45%, FACS analysis). By contrast, fasting cycles (without FBS) resulted in higher proliferation rate of *CROCC*-deficient cells than control (by more than 50% at 72h), suggesting an impaired cell cycle progression as a consequence of mitotic replication stress or misaligned chromosomes.[67,71] Most strikingly, *CROCC*-deficient cells exhibited all cardinal signs of rhabdoid features, displaying huge nuclei pushed to the periphery of the cells with single or multiple large nucleoli associated with eosinophilic cytoplasmic inclusions and large cellular protrusions resembling morphology observed *in vivo***.** **T**hese features resulted in dramatic changes of spindle-shaped morphology, activation of prometastatic genes[4,29] and enhanced metastatic potential which recapitulated the molecular phenotype seen in patients**(Fig 14).** Together,

these data show that rhabdoid phenotype underlie *CROCC* deficiency in *BRAF*-mutated colon cancer.

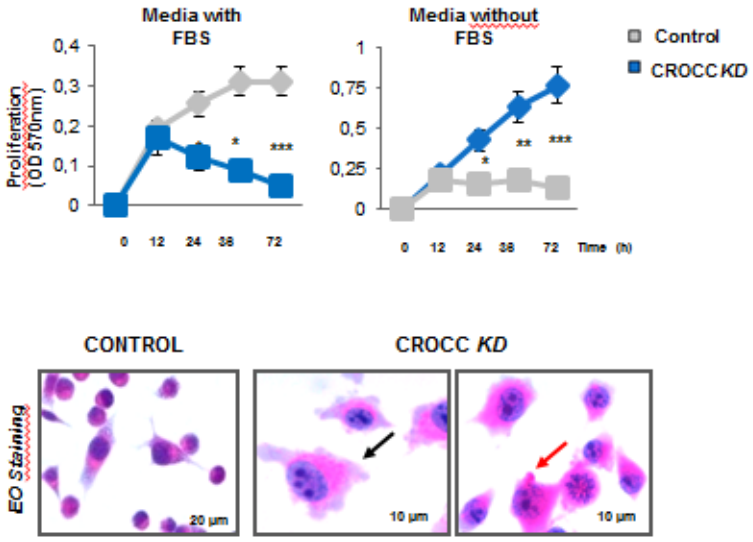


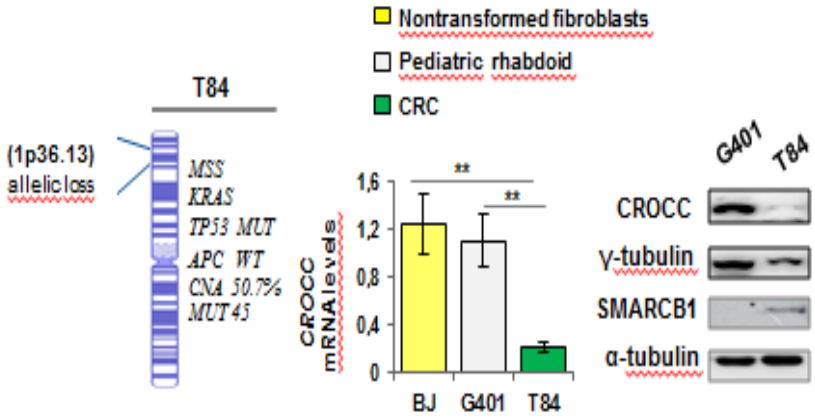
Fig 14 Proliferation of CROCC KD and control cells subjected to FBS (10%) supplementation or to fasting cycles "without FBS". The *P≤0.05, **P<0.01 and ***P<0.001 were derived from two-tailed Student's t-test. The graphs show mean values and standard deviation of five repeats. CROCC-depleted and control cells stained for hematoxylin and eosin (H&E); black arrow point to characteristic rhabdoid-like features with large polygonal cells and eccentric round nuclei with prominent nucleoli in CROCC KD cells. Red arrow point to typical eosinophilic hyaline cytoplasmic inclusions of malignant cells with rhabdoid-like features. Morphology of control cells are reported.

4.4 *CROCC* restoration suppresses growth in a metastatic model harboring 1p36.13 allelic loss

To test the hypothesis that *CROCC* impacts tumor growth and centrosome-related mitotic errors, we analyzed T84 CRC cells predict to exhibit copy-number loss at 1p36.13 locus and a high metastatic potential as being derived from a lung metastasis.[67] Indeed, we found recurrent "monopolar spindles", tetraploid or near-tetraploid cells, and an increased rate of micronuclei consistent with reduced *CROCC* endogenous activity, many of which were dramatically augmented in growth starvation cycles then overlapping *CROCC*-knockdown cells. In these latter conditions, immunofluorescence for CROCC revealed a proportion of cells either negative (40%) or displaying a single (30%) or fragmented centrosome (10%) in interphase not properly overlapping with γ-tubulin localization**.** Most strikingly, such aberrations were rarely, if ever, detected across pediatric rhabdoid G401, colon cancer retai50ning 1p36.13 locus or BJ cells**.** Therefore, we

transfected CROCC-*GFP*-tagged (1–2018aa) into T84 cells.[24] Unexpectedly, restoration of CROCC, showed a dramatic decrease of cell viability (12 days later, 0%) as compared to control plasmid, GFP alone. Similarly, the colony formation assay after 10 days, showed a robustly induced 7-fold fewer colonies than those transfected with a control (data not shown). Moreover, gain of *CROCC* conferred a flat/adherent phenotype and formation of filament-like structures co-localizing with γ-tubulin resulting in a profound actin cytoskeleton reorganization and lower expression of prometastatic and centrosome-related genes then control (**Figure 15 A,B**).[4]

A



B

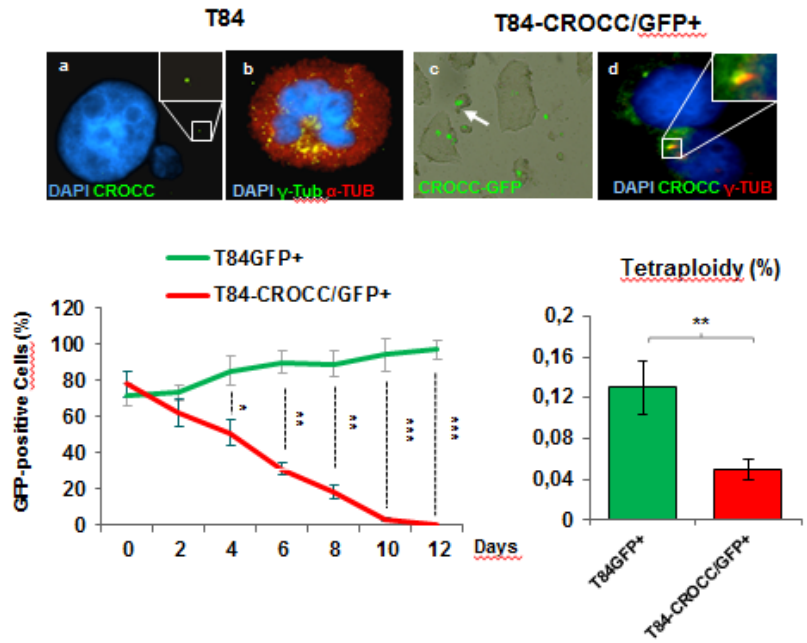


Figure 15. *CROCC* rescue corrects mitotic errors in a colorectal cancer harboring 1p36.13 allelic loss.

A) T84 colorectal cancer cells exhibit copy-number loss at 1p36.13 *CROCC* locus by *in silico* prediction. Deletion at the 1p36.13 locus induces loss of *CROCC* mRNA expression as compared to pediatric-rhabdoid-derived G401 cells or normal human cells, foreskin fibroblasts (BJ) with intact 1p36.13 locus. Percentage of segregation errors taking into account micronuclei and monopolar spindle defects, centrosome defects, chromosomal ploidy and mutated tumor suppressors in T84 and G401 cells (n>100 per cell line). **B)** Images of micronucleated cells stained for CROCC (green, enlarged in insets) and DNA (blue, dapi) in T84 basal cells. Scale bars, 10 μM. Example of monopolar spindle with fragmented centrosome stained for microtubules (α-tubulin), centrioles (γ-tubulin) and DNA (dapi). T84 cells transfected with wild-type CROCC-GFP or (GFP alone see supplementary material for details) and immunostained for γ-tubulin (red, enlarged in insets). Below on the left, cells transfected with wild-type *CROCC* and control are maintained in neomycin ($\mu g\ ml^{-1}$) for the indicated time. The *P≤0.05, **P<0.01 and ***P<0.001 were derived from two-tailed Student's t-test. Right, chromosome number detected in metaphases (after 5 days) are quantified by tetraploid on diploid cells ratio (22 chromosome spreads were analyzed in each condition). The **P<0.01 was derived from two-tailed Student's t-test.

5 Discussion

In this study, we have shown that genetic alterations in *CROCC* are related to a subset of rare rhabdoid cancers arising in the colon of adult patients, characterized by a lethal clinical course irrespective of stage at diagnosis. Although these tumors are morphologically indistinguishable from their pediatric counterparts, here we show that rhabdoid phenotype arise from distinctive pathways, supporting their epithelial origin. The high mutational load further illustrates the biologic distinction between RC and pediatric rhabdoid tumors, in which it is well-established the order of 1–10 mutations per exome reminding a genetic complexity similar to normal cells.[60,62,63] Notably, we found that RCs share mutation in genes that participate to centrosome and microtubule-dependent functions, suggesting an alternative route towards aneuploidy and chromosome instability. This is particularly relevant because exome data from ~5,000 human cancers indicate that genetic alterations in centrosomal genes are relatively rare.[1,62,63,41]

Our data unveil that impaired *CROCC* activity either caused by somatic mutations or copy number loss at the 1p36.13 locus "in a heterozygous state" result in defective centrosomes phenotypes without extra centrosomes, gross mitotic errors and increased tetraploidy**.** In contrast, related cancers lacking *CROCC* alterations "pediatric rhabdoid or prototypical colorectal tumors", display recurrent CROCC overexpression resulting in morphologically normal, larger or supernumerary centrosomes.[1-4,] How does CROCC facilitate mitotic errors? This gene is essential to establishing and maintaining interconnected centrosome after centrioles disengagement, and its phosphorylation by several kinases stimulates premitotic centrosome disjunction thereby ensuring mitotic spindle fidelity. In the RC cases, is conceivable that defective *CROCC* could compromise its controlled degradation by kinases impairing bipolar mitotic spindle assembly.[58] Indeed, *CROCC* knockdown in CRC cells lacking aneuploidy, causes structurally compromised centrosomes, monopolar spindle DNA segregation errors, increased formation of micronuclei and tetraploid cells. In this respect, tetraploid cells were characterized by stoichiometric double amount of DNA, a larger cell volume miming rhabdoid-like phenotype observed *in vivo*. Finally, *CROCC* restoration into cells harboring allele loss at the 1p36.13 locus, further confirmed its critical role as biological barrier against replication stress and tetraploidy, both common route to chromosomal instability and intratumour heterogeneity.[3,67] These findings highlight that genetic aberrations in critical centrosomal genes may be of broad general relevance to cancer or be exploited therapeutically by cytotoxic or microtubule destabilizing agents[72,73]. Therefore, in some contexts centrosome anomalies "amplification or deletion", could confer advantageous characteristics that promote tumor progression.[4,74,75] So far however, mutations in centrosome genes with a critical role

in procentriole formation, CEP57, *CEP135* and *PLK4* kinase, have been only identified in rare diseases, particularly microcephaly and Seckel syndrome.[71,74] Future experiments will aim to determine how centrosome anomalies are connected with pathways such as "*TP53*, WNT/β-catenin or Hippo" that play a key role in safeguarding the integrity of the human genome. In conclusion, the study of rare and lethal human tumors may allow molecular insight into crucial pathways that can have profound consequences on cancer cells reprogramming and shape.

References

1) Gönczy P. Centrosomes and cancer: revisiting a long-standing relationship. Nat Rev Cancer 2015;15:639-52.

2) Nigg, E. A. & Raff, J. W. Centrioles, centrosomes, and cilia in health and disease. Cell 139, 663–678 (2009)

3) Gordon, DJ, Resio, B, Pellman D. Causes and consequences of aneuploidy in cancer. Nature Rev Genet 2012;13:189–203.

4)Aparna Kalyan, Gurleen Pasricha, Dulabh Monga, Aatur Singhi, Nathan Bahary
Case Report of Rhabdoid Colon Cancer and Review of Literature Clinical Colorectal Cancer, Vol. 14, No. 1, e5-8 2015

5) Beckwith JB, Palmer NF. Histopathology and prognosis of Wilms' tumor: results from the First National Wilms' Tumor Study. Cancer 1978;41:1937-48.

6) Massimo Pancione, Arturo Di Blas, Lina Sabatino, Alessandra Fucci, Anna Maria Dalena, Natalia Palombi, Pietro Carotenuto, Gabriella Aquino, Bruno Daniele, Nicola Normanno, Vittorio Colantuoni. A novel case of rhabdoid colon carcinoma associated with a positive CpG island methylator phenotype and BRAF mutation Human Pathology (2011) 42, 1047–1052

7) Baba Y, Uchiyama T, Hamada K, Ishihara Y, Tanaka H, Isono Y, Saito T, Mukai K, Murata T, Ajioka Y
A case report of undifferentiated carcinoma of the sigmoid colon with rhabdoid features Nihon Shokakibyo Gakkai Zasshi. 2014 Jul;111(7):1384-90.

8) Romera Barba E, Sánchez Pérez A, Duque Pérez C, García Marcilla JA, Vázquez Rojas JL. Malignant rhabdoid tumor of the colon: a case report. Cir Esp. 2014 Nov;92(9):638-40

9) Chetty R[1], Bhathal PS. Caecal adenocarcinoma with rhabdoid phenotype: an immunohistochemical and ultrastructural analysis Virchows Arch A Pathol Anat Histopathol. 1993;422(2):179-82.

10) Yang AH[1], Chen WY, Chiang H. Malignant rhabdoid tumour of colon. Histopathology. 1994 Jan;24(1):89-91.

11) V. A. Marcus, J. Viloria, D. Owen, and M.-S. Tsao, "Malignant rhabdoid tumor of the colon. Report of a case with molecular analysis," Diseases of the Colon & Rectum, vol. 39, no. 11, pp. 1322–1326, 1996.

12) Nakamura, K. Nakano, K. Nakayama et al., "Malignant rhabdoid tumor of the colon: report of a case," Surgery Today, vol. 29, no. 10, pp. 1083–1087, 1999.

13) T. Kono, Y. Imai, J. Imura et al., “Cecal adenocarcinoma with prominent rhabdoid feature: report of a case with immunohistochemical, ultrastructural, and molecular analyses,” International Journal of Surgical Pathology, vol. 15, no. 4, pp. 414–420, 2007.
14) A. Mastoraki, O. Kotsilianou, I. S. Papanikolaou, P. G. Foukas, G. Sakorafas, and M. Safioleas, “Malignant rhabdoid tumor of the large intestine,” International Journal of Colorectal Disease, vol. 24, no. 11, pp. 1357–1358, 2009.

15) Seok Ju Lee, M.D., Tae Ho Kim, M.D., Dong Hoon Ko, M.D., Jeung Ah Kim, M.D., Chang Whan Kim, M.D., Jean A Kim, M.D.†, Do Sang Lee, M.D.* and Sok Won Han, M.D. Undifferentiated Adenocarcinoma of the Colon with Rhabdoid Features Korean Journal of Gastrointestinal Endoscopy 2010;40(1): 49-53.
16) Hoon Kyu Oh, Chang Ho Cho, Yoon Seup Kum Adenocarcinoma of the Sigmoid Colon with Prominent Rhabdoid Features: A Case Report. The Korean Journal of Pathology 2008;42(1): 63-65.
17) Sang Hwa Lee, Hyesil Seol, Wook Youn Kim, So Dug Lim, Wan Seop Kim, Tae Sook Hwang, Hye Seung Han Rhabdoid Colorectal Carcinomas: Reports of Two Cases The Korean Journal of Pathology 2013; 47: 372-377
18) A. Agaimy, T. T. Rau, A. Hartmann, and R. Stoehr, “SMARCB1 (INI1)-negative rhabdoid carcinomas of the gastrointestinal tract: clinicopathologic and molecular study of a highly aggressive variant with literature review,” American Journal of Surgical Pathology, vol. 38, no. 7, pp. 910–920, 2014.
19) A. Remo, C. Zanella, E. Molinari et al., “Rhabdoid carcinoma of the colon: a distinct entity with a very aggressive behavior: a case report associated with a polyposis coli and review of the literature,” International Journal of Surgical Pathology, vol. 20, no. 2, pp. 185–190, 2012.
20) Narimantas Evaldas Samalavicius, Rokas Stulpinas,[1] Valdas Gasilionis,[2] Edita Baltruskeviciene, Eduardas Aleknavicius, and Ugnius Mickys Rhabdoid Carcinoma of the Rectum Ann Coloproctol. 2013 Dec; 29(6): 252–255.
21) E. Moussaly and J. P. Atallah A Rare Case of Undifferentiated Carcinoma of the Colon with Rhabdoid Features: A Case Report and Review of the Literature Case Reports in Oncological Medicine
Volume 2015 (2015), Article ID 531348
22) H.-K. Oh, C.-H. Cho, and Y.-S. Kum, “Adenocarcinoma of the sigmoid colon with prominent rhabdoid features—a case report,” Korean Journal of Pathology, vol. 42, no. 1, pp. 63–65, 2008.
23) E. Romera Barba, A. Sánchez Pérez, C. Duque Pérez, J. A. García Marcilla, and J. L. Vázquez Rojas, “Malignant rhabdoid tumor of the colon: a case report,” Cirugia Espanola, vol. 92, no. 9, pp. 638–640, 2014.

24) Gadd S, Sredni ST, Huang CC, Perlman EJ. Rhabdoid tumor: gene expression clues to pathogenesis and potential therapeutic targets. Lab Invest 2010;90:724–738

25) SIMONE T. SREDNI AND TADANORI TOMITA Rhabdoid Tumor Predisposition Syndrome Pediatric and Developmental Pathology 18, 49–58, 2015

26) Biegel JA, Allen CS, Kawasaki K, Shimizu N, Budarf ML, Bell CJ. Narrowing the critical region for a rhabdoid tumor locus in 22q11. Genes Chromosomes Cancer 1996;16:94–105

27) Schneppenheim R, Fruhwald MC, Gesk S, et al. Germline nonsense mutation and somatic inactivation of SMARCA4/BRG1 in a family with rhabdoid tumor predisposition syndrome. Am J Hum Genet 2010;86:279–284

28) Bourdeaut F, Lequin D, Brugieres L, et al. Frequent hSNF5/INI1 germline mutations in patients with rhabdoid tumor. Clin Cancer Res Off J Am Assoc Cancer Res 2011;17:31–38

29) Pancione M, Remo A, Sabatino L, Zanella C, Votino C, Fucci A, Di Blasi A, Lepore G, Daniele B, Fenizia F, Molinari E, Normanno N, Manfrin E, Vendraminelli R, Colantuoni V. "Right-sided rhabdoid colorectal tumors might be related to the Serrated Pathway". Diagnostic Pathology. 2013 Feb 20; 8(1): 31.

30) Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics (Oxford, England)* 25 (14): 1754–60. doi:10.1093/bioinformatics/btp324.

31) Douville, Christopher, Hannah Carter, Rick Kim, Noushin Niknafs, Mark Diekhans, Peter D. Stenson, David N. Cooper, Michael Ryan, and Rachel Karchin. 2013. "CRAVAT: Cancer-Related Analysis of Variants Toolkit." *Bioinformatics* 29 (5): 647–48. doi:10.1093/bioinformatics/btt017.

32)DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran V Garimella, Jared R Maguire, Christopher Hartl, Anthony A Philippakis, et al. 2011. "A Framework for Variation Discovery and Genotyping Using next-Generation DNA Sequencing Data." *Nature Genetics* 43 (5): 491–98. doi:10.1038/ng.806.

33)Cibulskis K et al. "sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature Biotechnology 31 (3), 2013: 213-19

34) Wang K et al. "ANNOVAR": Functional Annotation of Genetic Variants from High-Throughput sequencing data. Nucleic Acids Research 38(16): 2010

35) Van Loo P, Nordgard SH, Lingjaerde OC, Russnes HG, and Rye IH, et al (2010). Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A 107, 16910–16915.

36) Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. Genome Res. 19, 1639–1645 (2009).

37) Roberta Spinelli et al. Identification of novel point mutations in splicing sites integrating whole-exome and RNA-seq data in myeloproliferative diseases. Molecular Genetics & Genomic Medicine 2013; 1(4): 246–259

38)Schwarz, J. M., C. Rodelsperger, M. Schuelke, and D. Seelow. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. Nat. Methods 7:575–576.

39)Skandalis, A., M. Frampton, J. Seger, and M. H. Richards. 2010. The adaptive significance of unproductive alternative splicing in primates. RNA 16:2014–2022.

40) Xing Hua et al. DrGaP: A Powerful Tool for Identifying Driver Genes and Pathways in Cancer Sequencing Studies. The American Journal of Human Genetics 93, 439–451, September 5, 2013

41) Cancer Genome Atlas Network. Comprehensive Molecular Characterization of Human Colon and Rectal Cancer.Nature 487: 330–37, 2012.

42) Sadanandam A, Lyssiotis CA, Homicsko Ket al:A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nature Med19(5): 619-25, 2013

43) Li J, Lupat R, Amarasinghe KC, Thompson ER, Doyle MA, Ryland GL, Tothill RW, Halgamuge SK, Campbell IG, Gorringe KL. CONTRA: copy number analysis for targeted resequencing. Bioinformatics. 2012 May 15;28(10):1307-13.

44)Barretina J, Caponigro G, Stransky N et al: The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483(7391):603-7, 2012.

45)Mouradov D, Sloggett C, Jorissen RN, Love CG, Li S, Burgess AW, Arango D, Strausberg RL, Buchanan D, Wormald S, O'Connor L, Wilding JL, Bicknell D, et al. Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. Cancer Res. 2014;74:3238–47.

46) William Stafford Noble. How does multiple testing correction work? Nat Biotechnol. 2009 Dec; 27(12): 1135–1137.

47)Zamo A, Bertolaso A, van Raaij AW, et al: Application of microfluidic technology to the BIOMED-2 protocol for detection of B-cell clonality. J Mol Diagn 14:30-7, 2012

48)Simbolo M, Gottardi M, Corbo V, et al: DNA qualification workflow for next generation sequencing of histopathological samples. PLoS One 8:e62692, 2013

49)Cingolani P, Patel VM, Coon M, et al: Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet 3:35, 2012

50)McLaren W, Pritchard B, Rios D, et al: Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26:2069-70, 2010

51)Robinson JT, Thorvaldsdottir H, Winckler W, et al: Integrative genomics viewer. Nat Biotechnol 29:24-6, 2011

52)Cho, E.H.; Whipple, R.A.; Matrone, M.A.; Balzer, E.M.; Martin, S.S. Delocalization of gamma-tubulin due to increased solubility in human breast cancer cell lines. *Cancer Biol. Ther.* 2010, *9*, 66–76.

53)Teixido-Travesa, N.; Roig, J.; Luders, J. The where, when and how of microtubule nucleation—One ring to rule them all. *J. Cell Sci.* 2012, *125*, 4445–4456.

54) Pancione M, Remo A, Zanella C, Sabatino L, Di Blasi A, Laudanna C, Astati L, Rocco M, Bifano D, Piacentini P, Pavan L, Purgato A, Greco F, Talamini A, Bonetti A, Ceccarelli M, Vendraminelli R, Manfrin E, Colantuoni V."The chromatin remodelling component SMARCB1/INI1 influences the metastatic behavior of colorectal cancer through a gene signature mapping to chromosome 22." Journal of Transational Medicine. 2013 Nov 28;11:297. doi: 10.1186/1479-5876-11-297.

55) Caracciolo V, D'Agostino L, Dráberová E, Sládková V, Crozier-Fitzgerald C, Agamanolis DP, de Chadarévian JP, Legido A, Giordano A, Dráber P, Katsetos CD. Differential expression and cellular distribution of gamma-tubulin and betaIII-tubulin in medulloblastomas and human medulloblastoma cell lines. J Cell Physiol. 2010 May;223(2):519-29.

56) Jackson EM, Shaikh TH, Gururangan S, et al. High-density single nucleotide polymorphism array analysis in patients with germline deletions of 22q11.2 and malignant rhabdoid tumor. Hum Genet. 2007;122:117–127.

57) J. L. Montgomery, L. N. Sanford, C. T. Wittwer: High-resolution DNA melting analysis in clinical research and diagnostics. *Expert Rev Mol Diagn* 10(2), 219-240 (2010)

58) Bahe S, Stierhof YD, Wilkinson CJ, Leiss F, Nigg, EA. Rootletin forms centriole-associated filaments and functions in centrosome cohesion. J. Cell Biol 2005; 171:27–33.

59) Votino C, Laudanna C, Parcesepe P, Giordano G, Remo A, Manfrin E, Pancione M. Aberrant BLM cytoplasmic expression associates with DNA damage stress and hypersensitivity to DNA-damaging agents in colorectal cancer. J Gastroenterol. 2016 May 11.

60)Lee RS, Stewart C, Carter SL. A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. J Clin Invest 2012; 122:2983-8.

61) Kandoth, C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333–339.

62) Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.

63) Shirley BC, Mucaki EJ, Whitehead T, Costea PI, Akan P, Rogan PK. Interpretation, Stratification and Evidence for Sequence Variants Affecting mRNA Splicing in Complete Human Genome Sequences. Genomics Proteomics Bioinformatics 2013;11:77–85.

64) Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603-7.

65)Henrich KO, Schwab M, Westermann F. 1p36 tumor suppression—a matter of dosage? Cancer Res 2012;72:6079–88.

66)Mouradov D, Sloggett C, Jorissen RN, et al. Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. Cancer Res 2014;74:3238–47.

67)Burrell RA, McClelland SE, Endesfelder D, et al. Replication stress links structural and numerical cancer chromosomal instability. Nature 2013;494:492-6.

68)Weischenfeldt J, Symmons O, Spitz F, Korbel JO. Phenotypic impact of genomic structural variation: insights from and for human disease. Nature Rev Genet 2013;14:125–138.

69) Conroy PC, Saladino C, Dantas TJ, Lalor P, Dockery P, Morrison CG.C-NAP1 and rootletin restrain DNA damage-induced centriole splitting and facilitate ciliogenesis. Cell Cycle 2012;11:3769-78.

70) Ghadimi, BM, Sackett DL, Difilippantonio MJ, et al. Centrosome amplification and instability occurs exclusively in aneuploid, but not in diploid colorectal cancer cell lines, and correlates with numerical chromosomal aberrations. Genes Chromosomes Cancer 2000;27:183–90.

71) Kalay E, Yigit G, Aslan Y, et al. CEP152 is a genome maintenance protein disrupted in Seckel syndrome. Nat Genet 2011;43:23–6.

72)Horazdovsky R, Manivel JC, Cheng EY. Surgery and actinomycin improve survival in malignant rhabdoid tumor. Sarcoma. 2013;2013:315170.

73)L. Vecchione, V. Gambino, J. Raaijmakers, et al., A vulnerability of a subset of colon cancers with potential clinical utility, Cell 165 (2016) 317e330.

74)Martin CA, Ahmad I, Klingseisen A, et al. Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. Nat Genet 2014;46:1283–92.

75)Basto, R., Brunk, K., Vinadogrova, et al. Centrosome amplification can initiate tumorigenesis in flies. Cell 2008; 133:1032-42.